Vol. VI Part 1 February 1954

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Editorial Board

D. RUSSELL DAVIS
R. C. OLDFIELD J. W. WHITFIELD
G. C. DREW M. D. VERNON
G. C. GRINDLEY O. L. ZANGWILL

CONTENTS

D. E. BROADBENT. Some Effects of Noise on Visual Performance - - - - - - - - 1

J. A. DEUTSCH. A Machine with Insight - - 6

L. C. T. JONES. Frustration and Stereotyped Behaviour in Human Subjects - - - - - - - - 12

M. WERTHEIMER. Can a Trained Subject Judge his Auditory Sensitivity? - - - - - - 21

R. P. KELVIN. Discrimination of Size by Sight and Touch - - - - - - - - - - - 23

N. S. SUTHERLAND. Figural After-Effects, Retinal Size, and Apparent Size - - - - - - - 35

Proceedings of Experimental Psychology Group - - 45

List of Members - - - - - - - - - 45

PUBLISHED FOR
THE EXPERIMENTAL PSYCHOLOGY GROUP
BY
W. HEFFER & SONS LTD. CAMBRIDGE

OCCUPATIONAL PSYCHOLOGY

Editor: ALEC RODGER

October, 1953 Volume XXVII, No. 4

CONTENTS

Difficulties of Older People in Industry. By R. M. BELBIN

Attitudes of Women Operatives to Promotion. By ELIZABETH LIVINGSTONE

Some Effects of Joint Consultation on the Status and Role of the Supervisor. By H. CAMPBELL

Seasonal Fluctuation in the Average Ability of Men Available for National Service. By SYLVIA ANTHONY

The Mobility of Labour in Great Britain. By GEOFFREY THOMAS

Book Reviews

Other Books Received

Index

Annual Subscription 30*s.*

NATIONAL INSTITUTE OF INDUSTRIAL PSYCHOLOGY
14, WELBECK STREET, LONDON, W.1.

THE BRITISH JOURNAL OF PSYCHOLOGY

(GENERAL SECTION)

Edited by D. W. Harding

Vol. XLV Part 1 **February, 1954** **12s. 6d. net**

L. S. HEARNSHAW. The psychological study of conceptual thinking.

HILARY M. CLAY. Changes of performance with age on similar tasks of varying complexity.

R. C. OLDFIELD. Memory mechanisms and the theory of schemata.

C. B. GIBBS. The continuous regulation of skilled response by kinaesthetic feed back.

D. SHEPPARD. The adequacy of everyday quantitative expressions as measurements of qualities.

E. C. POULTON. Verbal extrapolation and interpolation.

M. B. SHAPIRO. A preliminary investigation of the effects of continuous stimulation on the perception of 'apparent motion'.

CORRESPONDENCE.

PUBLICATIONS RECENTLY RECEIVED.

The subscription price per volume, payable in advance, is 40s. net (*post free*).

Subscriptions may be sent to any bookseller or to the

CAMBRIDGE UNIVERSITY PRESS

Bentley House, Euston Road, London, N.W.1

1127

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. VI 1954 Part 1

SOME EFFECTS OF NOISE ON VISUAL PERFORMANCE

BY

D. E. BROADBENT

From the Medical Research Council Applied Psychology Research Unit, Cambridge.

A group of ten subjects showed impaired performance, when watch-keeping on a display made up of steam-pressure gauges, in 100 db. noise as compared with 70 db. On the easier task of watch-keeping on a display made up of small lights, another group of twenty subjects showed no overall effect of noise. Individual subjects who showed a practice effect on the latter task comparable to that shown by all subjects on the former one, however, also showed a similar effect of noise.

In addition, performance on the light-watching became relatively less efficient in noise with continued exposure: and although parts of the task were still adequately carried out, others were not. The fact that noise effects are thus functions of individual differences, of visibility of signal, and of length of performance in noise, allows us to explain the negative findings of many previous workers.

I

Introduction

Recent reviews of the work on the psychological effects of noise (Berrien, 1946; Kryter, 1950, 1952) state correctly that there is no adequate existing evidence of any effect beyond the initial disturbance produced by startling noises. Yet as Konorski (1948) points out, extremely valuable information about the nature of processes in the nervous system is likely to be derived from interference and facilitation between them. If those who complain of effects of noise have objective grounds for doing so, it may be that theories of behaviour should consider the rôle of background stimulation more thoroughly than they do. The factual issue is thus of some importance, and the following experiments were designed to cover certain possibilities which the existing work still left open. These possibilities may be summarized thus:

1. The effect of task difficulty. It has been shown by Mackworth (1950, p. 57) that "fatigue" effects are more easily shown with faint signals than with easily visible ones, and a similar relation may apply to other forms of stress.

2. The rôle of individual differences. It has been found by the writer in preliminary experiments that "fatigue" effects appear more readily in certain individuals, who may be detected by the size of the practice effect between runs. A similar relation may apply to other forms of stress.

3. The effect of prolonged performance in noise. It is in general true (Babington-Smith, 1951) that frequent changes in conditions do not allow human performance to settle to a level characteristic of each condition: prolonged performance in noise may therefore be needed to show decrement.

4. The use of analytic rather than overall scores. Complex tasks may give similar overall scores under stress although the relation between parts of the performance may in fact be changing: a point put for instance by Bartlett (1945).

II

Procedure

Production of noise. A room was set up sixteen feet square and eight feet high, built of metal plate loaded to prevent resonance. In three of the walls loudspeakers were inserted, each wall carrying one H/F speaker with a multi-cell horn and one L/F speaker with re-entrant horn. The frequency of cross-over was 1000 cycles. The speakers were fed from an R.G.D. tape recorder through 50-watt Vortexion amplifiers, one for each pair of speakers. This arrangement gave a flat field over the whole area in which the subject worked.

The noise itself was recorded on an endless loop of tape, the recording having been made near rapidly moving machinery. The spectrum was reasonably flat ($\pm$ 3 db.) on third-octave analysis from 40–5000 c.p.s. On "noise" days the level used was 100 db. re 0·0002 dynes/sq. cm., and on "quiet" days the same recording was played at 70 db. to mask minor extraneous noises. The levels were checked daily by a General Radio sound-level meter which was itself checked at intervals against Admiralty Research Laboratory equipment.

The tasks employed. Two visual vigilance tasks were given to separate groups of subjects. The first, the Twenty Dials Test, required the subject to watch twenty steam-pressure gauges: and if he saw any of the pointers reading above a danger mark, to turn a knob below the gauge to bring the pointer back below the mark. The gauges were arranged on three sides of a square of 12 ft. side, the subject being seated in the middle of the open side. Each gauge was six inches in diameter; the pointers were stationary except when a signal was delivered, when the tip of the pointer moved through 0·8 in to a new position above the danger mark and again became stationary. A silent mechanism operating on the principle of the ammeter was employed to move the pointer. The length of watch was one and a half hours, and fifteen signals were delivered in this time at intervals varying from 1 to 12 minutes. Each half-hour of the watch contained a signal at each of the intervals used, and position of the signal in the display was similarly counter-balanced.

In various preliminary experiments on sixty subjects, it was found that the first watch on this task was anomalous due to practice, but that when it is repeated five times at twenty-four hour intervals under quiet conditions the last four watches do not vary significantly in performance. If anything, performance is better on the third and fourth days. Consequently for this investigation ten subjects were given the task five times, with "noise" on the third and fourth days and the remaining days "quiet."

The second task, the Twenty Lights Test, was similar to the Twenty Dials except that each dial was replaced by a head-lamp bulb run at half-voltage and covered with paper so as to give a dim glow. Each bulb could be clearly recognised as "on" or "off" in foveal vision, but did not attract attention in the periphery: their brightness was approximately 5 foot-lamberts against a background of 3 foot-lamberts. To prevent the task becoming too easy, the Lights were not all at eye-level as the Dials were, but irregularly arranged from ground-level to a height of four feet. The response in this case was to press a key when a lamp was seen alight, the distance of the key from the subject's chair being the same as that of the knobs in the Dials Test. The other details of administration were the same as for the Dials.

Ten subjects were given this task five times each, with noise on the third and fourth days: since this task had not been used repeatedly in quiet like the Dials, a further ten subjects were given the task five times, with noise on the first and second days.

Both on the Dials and Lights, the experimenter and controlling apparatus were in a control room separated from the metal chamber by an air space and a brick wall. As a check on wakefulness, etc., the subject could be observed from time to time without his knowledge. All subjects were Naval Ratings aged 18–30 years. It is important to note that almost all of them were familiar with noise levels at least as high as the one used from their ordinary work, but that this did not prevent effects from appearing. Twenty of them were Fleet Air Arm ratings (mostly mechanics), three were stokers, four seamen, and three radar operators.

III

Results

Effect of task difficulty. Responses on the Dials are divided into "seens" and "founds," the former being those in which the subject reports that he was looking at the dial at the instant when the pointer moved, and so had his attention caught by

his conspicuous signal. These show no significant effect of noise. In the "founds," he proportion responded to in nine seconds or less is taken as the "quick found" core. The mean proportion of "quick founds" for each run is shown in Table I. As in the earlier work, there is a practice effect between runs 1 and 2. The difference etween runs 3 and 4 in noise and runs 2 and 5 in quiet is significant, t = 3·21 while he 0·05 level is 2·26 and the 0·01 level is 3·25.

On the Lights, the proportion of "quick founds" is higher throughout, the signal eing a much more conspicuous one. On the group given the same testing schedule s the Dials group, there is no practice effect between runs 1 and 2. Nor is there ny difference between noise and quiet in either of the two Lights groups. The more ifficult signal, therefore, shows an effect of noise while the easier one does not.

TABLE I

THE PROPORTION OF "QUICK FOUNDS" IN NOISE AND QUIET

Task and Group	*Run*				
	1 *Quiet (Practice)*	2 *Quiet*	3 *Noise*	4 *Noise*	5 *Quiet*
Dials	0·215	0·339	0·193	0·244	0·381
Lights: the markedly improving subjects ..	0·424	0·574	0·328	0·374	0·454
Lights: the least improving subjects..	0·432	0·226	0·482	0·300	0·414

NOTE:—On the Lights Test the subjects shown are those treated similarly to the Dials subjects, so as to facilitate comparison. See text for the control of order effects by ther groups. The "markedly improving" subjects are the five showing the greatest alue of (Run 2—Run 1) and the "least improving" the five remaining.

Effect of individual differences. Despite the absence of any overall effect on the ights, individual differences appear which can be best illustrated comparatively y the method of Table I. Here the ten subjects treated similarly to the Dials roup are split into the five showing the greatest and the five showing the least nprovement from run 1 to 2. The effect for the former is strikingly parallel to hat for the Dials, while the latter behave quite differently. To test the significance f this difference, we may consider all twenty subjects tested on the Lights, eglecting the last quiet run (which was given to ensure that all subjects had milar expectancies). Each subject then has two successive runs under each ondition, half the subjects having noise first and half quiet. If we then correlate im-rovement between the two quiet runs with the amount of deterioration in noise, i.e.

$$(Q_2 - Q_1) \text{ and } (Q_2 + Q_1) - (N_2 + N_1)$$

e obtain a tau summed for the two groups of 0·36, which is significant, S/S.E. of = 2·09, while the 0·05 level is 1·96. This correlation is not due to exaggerated ndom variation in Q_2: both $(Q_2 - N_2)$ and $(Q_1 - N_1)$ correlate to approximately e same extent with $(Q_2 - Q_1)$. An independent test has also been used to differ-ntiate the susceptible individuals: see Discussion.

Effect of prolonged performance. In Table I we see the difference between quiet and noise performance on the Lights for each successive half-hour from the beginning of a particular condition. It will be noted that a slight initial decrement in noise is followed by an actual advantage for noise, which decreases progressively during the remaining half-hours. It is not possible to obtain an adequate estimate of error for each of these differences, since this experimental design (like any other) must confound certain interactions involving subjects which are known from the previous work to be important. By taking the scores for whole runs rather than half-hours, however, most of these are eliminated, and it then appears that $N_1 - N_2$ is very significantly greater than $Q_1 - Q_2$; that is, the general downward trend of Table I is confirmed. $F = 8{\cdot}46$ while the 0·01 level is 8·28. Performance in noise becomes worse relative to that in quiet with increasing length of exposure. It may be noted that this fact is consistent with the interpretation that the apparent absence of overall effect on the Lights is due to a compensatory effort by some subjects, and that this effort cannot be maintained.

TABLE II

DIFFERENCE BETWEEN NOISE AND QUIET ON THE LIGHTS

Half-hours since beginning of condition	1	2	3	Gap between runs	4	5	6
Noise-quiet (proportion of quick founds)	−0·046	+0·140	+0·058	—	−0·004	−0·101	−0·133

Use of analytic scores. This downward movement does not reach a point at which performance on the Lights is significantly worse in noise than in quiet, using the score so far employed. But certain component scores do show a change in the pattern of performance in noise. Thus, if we take the longest succession of signals each of which took more than 9 seconds to elicit a response, we find that the series is longer in the second noise run than in the second quiet run. Using a logarithmic transformation, $F = 5{\cdot}18$ while the 0·05 level is 4·41. The length of the sequence of misses corresponding to the mean of the transformed series is three signals for quiet and four for noise. Missed signals appear therefore to be more concentrated by noise into certain periods of the watch.

There is also a tendency for the lights in front of the subject to be neglected as compared with those at the sides. Considering the number of signals from the lights facing the subject responded to in 4 seconds or less, there is a mean advantage for the second quiet run as opposed to the second noise run of 0·6 signals, for which $t = 2{\cdot}45$ while the 0·05 level is 2·09. Similar responses to the lights on each side of the subject showed a slight improvement in noise, but quite insignificant ($t = 1{\cdot}28$)

Thus it appears that parts of the performance both in space and time are impaired by noise while other parts are not; so that overall scores show no significant effect, being diluted by the unaffected parts.

IV

DISCUSSION

In certain respects (task difficulty and the use of analytic scores) the present results show reasons which might explain the negative findings of the most careful previous workers: but these factors may or may not have been important in earlier

experiments. Other points, however, were quite certainly operative. Most important of these are the individual differences.

The individuals most susceptible to noise show the largest practice effects between the first two runs in quiet: and it may be noted that in the preliminary experiments it was found that this score also identified the subjects showing the largest "fatigue" effects. An independent test was found to correlate with fatigability both on the Dials and Lights, and also to show a relation to the effect of noise. This test was unrelated to test intelligence but showed significant links with academic attainment. It was a modification of the Level of Aspiration test used by Eysenck and his collaborators (1947) to distinguish dysthymics and hysterics, a dimension which is known to be related to education. It seems likely, then, that a sample of subjects selected by academic examination or similar means will also by that very fact be selected for resistance to the effects of continuous performance and of irrelevant attention-demanding stimulation. The present research was more fortunate than almost all its predecessors in having available a sample not drawn from the undergraduate or research worker population.

Secondly, the length of performance used in these experiments differs from earlier work. On the one hand tasks as long or longer have been performed with occasional short bursts of noise, and on the other hand very long exposures to noise have been investigated by short tests. But it is unusual to find previous work of comparable scope on the performance of one task in one condition of noise or quiet for periods as long as those used here. Yet, as the data of Table II show, this is clearly likely to falsify results. When a prolonged task and a long exposure were used by Davis (1948), the results suggested a real but complex effect of noise with marked individual differences.

Finally, and perhaps most important of these factors, is the nature of vigilance tasks themselves. It is likely that tasks requiring reception of information from a particular source at an uncertain time are peculiarly vulnerable to the effects of intense competing stimuli such as noise. This peculiarity of vigilance tasks requires further experimental evidence for its elucidation, and has been examined elsewhere (Broadbent, 1953). We may note here simply that earlier work on noise has not used this type of task.

This work was supported financially by the Medical Research Council, and thanks are due to the Royal Navy both for supplying subjects and for providing the very extensive technical facilities required. The influence of Professor Sir Frederic Bartlett, F.R.S., and of Dr. N. H. Mackworth will be readily apparent throughout the paper.

References

Babington-Smith, B. (1951). On some difficulties encountered in the use of factorial designs and analysis of variance in psychological experiments. *Brit. J. Psychol.*, **42**, 250–268.

Bartlett, F. C. (1945). Some growing points in experimental psychology. *Endeavour*, **4**, No. 14.

Berrien, F. K. (1946). The effects of noise. *Psychol. Bull.*, **43**, 141–161.

Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, **44**, 295–303.

Davis, D. R. (1948). Pilot Error. A.P. 3139a. London: H.M. Stationery Office.

Eysenck, H. J. (1947). *Dimensions of personality*. London.

Konorski, J. (1948). *Conditioned reflexes and neuron organisation*. Cambridge.

Kryter, K. D. (1950). The effects of noise on man. *J. Speech Dis., Monogr. Suppl.* No. 1.

Kryter, K. D. (1952). Noise safety criteria. *Arch. industr. Hyg.*, **5**, 117–120.

Mackworth, N. H. (1950). *Researches on the Measurement of Human Performance*. Medical Research Council Special Report No. 268. London: H.M. Stationery Office.

A MACHINE WITH INSIGHT

BY

J. A. DEUTSCH

From the Institute of Experimental Psychology, Oxford

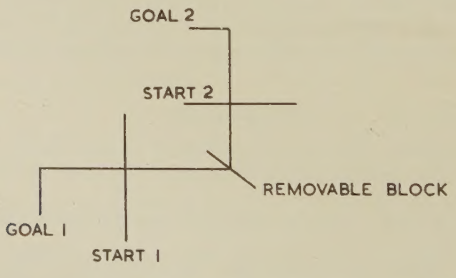
When an animal solves a novel problem without trial and error, a psychologist tends to call the behaviour insightful. A relatively simple machine resting on new principles, capable of this and learning, has been constructed by the writer in order to demonstrate his tentative explanation of this and other aspects of animal behaviour.

The machine consists of three parts. The first retains information and utilises it in accordance with the goal set. The second is a trolley which is guided by a pulse which the first part transmits; it also has bumpers which cause it reflexly to steer away from any obstacle and to turn out of corners by reversing one of the motors which drive each of the two main wheels separately; and it has a light mounted on it which is thrown forward at a wide angle. The third part comprises photocells, which act as receptors and are attached to the walls of the maze which the trolley is required to learn.

Behaviour

The machine's performance is more precisely described as follows. It learns and retains any two short, modified, rat mazes. On the learning run the machine is made to enter all the blind alleys of the maze on its way to the goal box. On the next run it will travel straight from the entrance of the maze to the goal box without entering a blind alley. If the two mazes, which it has learnt and retained, share a common point (even if this is down blind alleys in each maze), the machine will, on being set in an entrance, find its way without any further trial and error to whichever of the two goal boxes the demonstrator makes it seek. The machine appears to have this property in common with the rat, which suddenly integrates its past experience in accordance with its aim in the latent learning situation when a goal object is introduced.

FIGURE 1



After learning two mazes separately, the machine will solve the problem of finding its way from the entrance of one maze to the goal of the other.

Furthermore, the machine can transfer or generalize its "knowledge" to mazes of completely different shape and similar only in a highly abstract way. Each alley in the maze is marked by a different signal which is supplied to the machine. If the sequential order of the signals down the alleys with respect to each other is kept the same, even though the shape of the maze is quite altered, the machine can still find its way. It will execute an errorless run even though it may now have to

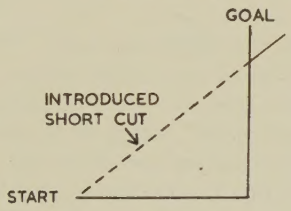
turn in opposite directions at the choice points. The machine does not learn a sequence of responses to a chain of appropriate stimuli. The movements that it makes are of little consequence both in its learning and its performance are always highly variable. It can be "taught" though the trolley be motionless and the machine be "paralysed" and make no motor response. The trolley will make an errorless run even when it is partially crippled.

FIGURE 2

The machine transfers its training from A to B and vice versa, but not from A to C.

The flexibility and variability of its performance may be illustrated by its ability to take advantage of short cuts.

FIGURE 3



The machine will take a short cut if it is made available.

After learning a maze, the machine will take a short cut, if one is introduced leading to a point nearer the goal. Similarly, it will take short cuts introduced when it is finding its way from the entrance of one maze to the goal in a second.

The machine can also be trained in operant conditioning situations, multiple choice discriminations and related problems.

The system

The behaviour described above is the behaviour of a certain type of system. The properties described are not those of relays but of arrangements of relays. Further, we can make the same arrangement, manifesting the same behaviour, out of other components—thermionic valves, transistors or wheels. Perhaps a similar arrangement is to be found in the mammalian central nervous system. After all, the mammal manifests similar behaviour.

The abstract system itself which is richer is described elsewhere (Deutsch, 1953). Here are described the operations which the machine performs and its particular embodiment as represented by this machine.

On the learning trial (before the goal is found) the system arranges the receptors in an order inverse to that in which they were stimulated. The receptor which was set off first will be last in the series, and the last one next to the first. The first position is always occupied by the goal-signalling receptor. There will therefore be a record of the order in which certain aspects of the environment occur in relation to the goal.

This order is preserved by making a pulse from a receptor travel to the trolley through a contact on relay with what might be called a particular number. The relay with the lowest number has the goal-signalling receptor already attached to it, and when it is switched on and closed by the operator, all the relays with higher numbers are closed, enabling any receptor attached to any of them to control the trolley, whichever it is that may be stimulated.

But when a pulse passes through any relay, all those relays which have a higher number will open and stay open, so that the receptors attached to them will not be able to influence the movements of the trolley. This will be called the hierarchical arrangement. It ensures that the trolley will always steer towards that "landmark" which it "saw" nearest (in time) to the goal.

When there are two series of receptors made, each registering the order of "cues" leading to the two goals, if there is a receptor common to both series, the system will exploit this information. When one of the goal-signal receptors is closed by the operator, it will be recalled that all the relays with higher numbers will also be switched on. Now one of the relays with a higher number has a receptor common to both series attached to it; through this common receptor will now flow excitation closing all the relays in the other series with a higher number. This gives the system "insight."

The machine

The system is embodied in the following way. First is the "thinking" part, composed of six uniselectors and twenty-three relays. The second is the trolley which moves through the maze, and the third consists of the photocell-receptors. The trolley has a light mounted on it, which is thrown forward at a wide angle, and which touches off the photocells mounted on each sector of the maze. (This arrangement was adopted as a practically easier equivalent of mounting receptors on the trolley and placing the signals in the maze.) The trolley has two responses: it either goes forward or turns round on its axis. When it has been set to learn its way to a goal, it goes forward (except when it bumps into the walls of the maze). When it has been set to find its way after it has learnt, it will continue to travel forward as long as a particular receptor which the machine has selected is being stimulated. When this stimulation ceases, the machine will turn until the stimulation of this receptor occurs again. It will thus guide itself towards the selected receptor. This is the way in which the trolley is controlled.

The way in which the machine stores the order in which the receptors were stimulated on the learning trial is now described. (See Figure 4.)

Each receptor is connected to its uniselector before it has been stimulated, in such a way that its stimulation will move the wipers of its own uniselector to the first position and the wipers of any other uniselector whose receptor has already been stimulated to the next position. Once it has been stimulated and the wipers shifted to the first position, the stimulation from its own receptor will not affect it; nor will it affect any other uniselector. It will now only be shifted by a receptor which is being stimulated for the first time.

This will be made clear if we study Figure 4. The uniselector U1, which is moved when receptor 1 is stimulated, has two principal states.

1. The state before receptor 1 has been stimulated. Here only the closing of relay R1, caused by the stimulation of receptor 1, will move the wipers of U1 from position O. Any pulse from R2, closed by the stimulation of receptor 2, will not affect U1 because the circuit which effects this is broken at U1, row b.

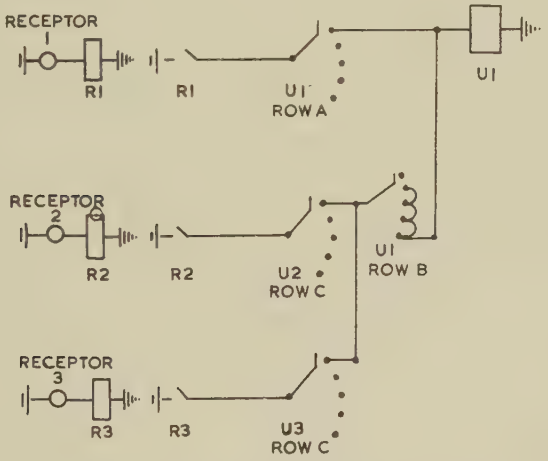
2. The state after receptor 1 has been stimulated. Here the closing of R1 will not further affect U1, because the wipers have now been shifted to position 1. And thus the circuit which shifts U1 when R1 closes will have been broken.

On the other hand the circuit which connects U1 with R2 and R3 is now closed by U1, row b. U1 will therefore be stepped on when R2 or R3 close. But U1 will only be affected if R2 (or R3) close for the first time because the circuit leading through U2 row c (or U3, row c) will be broken as soon as the wipers have been moved from position 0.

When considering receptor n, substitute n for 1 in the above description and vice versa.

Now, if there are three receptors stimulated in turn, it follows that the uniselector attached to the first one to be stimulated will be on the third position, the next on the second position and the last on the first position.

FIGURE 4



The circuit used for attaching receptor 1 in its place during learning.

For receptor 2, exactly the same circuit is used, but 2 should be read in place of 1, and vice versa.

The receptors are also connected through wipers to the positions, each of which is separately attached to the hierarchical arrangement. Thus the nearer in time to the goal a receptor was touched off, the higher in the hierarchy is it connected. Once the goal the machine was set is found, the connection becomes permanent. As on the learning run the trolley is made to "explore" each alley or sector in turn before it reaches the goal, the receptors are connected to the hierarchical arrangement in accordance with their proximity to the goal-receptor.

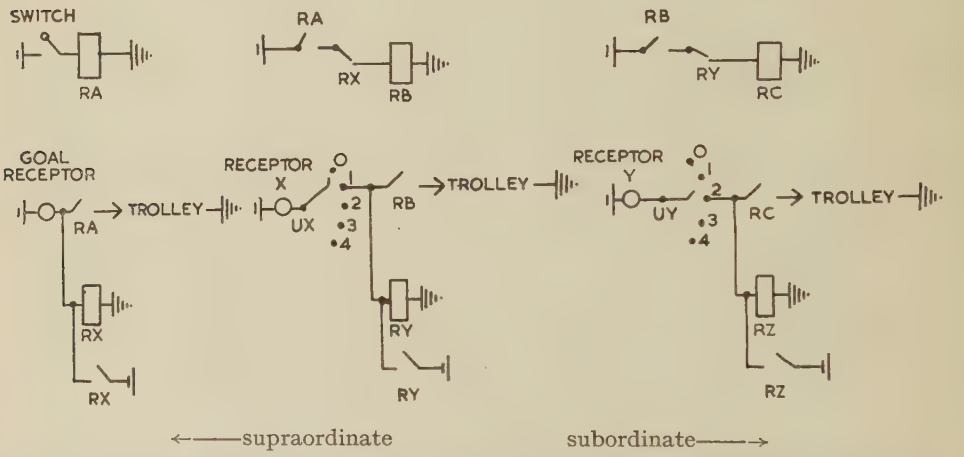
The order thus preserved is utilised in the following manner. (See Figure 5).

The receptors are connected by uniselectors to a hierarchical arrangement of relays. By this is meant that the relays are connected in a series in which the energizing of a preceding (supraordinate) relay causes the closing of its successor (subordinate), and the switching off a predecessor leads to the opening of the relay whose closing it previously caused.

The receptor whose stimulation functions as goal (two in the present model) is "innately" connected to the "make" contacts of the first relay in this series (RA) (again two in the present model) (see Figure 5). The rest of the receptors (x, y) are also connected to the "make" contacts of relays but each is connected to its respective relay in the hierarchy through the activity of its uniselector (Ux, Uy) during the "training" trial. Each receptor is connected at the same time to close another relay (RX, RY, RZ) which breaks the circuit of the relay subordinate to the one to which the receptor is also connected by the uniselector. When the hierarchy is switched on ("when there is a drive acting on the organism") the stimulation of any receptor connected to the hierarchy will cause the trolley to go forward while this stimulation continues, because it is connected to the trolley through the "make" contact on its relay in the hierarchy and secondly, will immediately lead to the opening, for the rest of the trial, of all the relays subordinate to that to which it is connected.

During the "training" trial the hierarchy is not switched on. Instead the circuit which allows the receptors to drive the uniselectors is closed, to be opened when the finding of the goal receptor is signalled.

FIGURE 5



The so-called hierarchical arrangement which selects the "cue" which the trolley is to approach. The make contact of relay RB is connected to the first position of every uniselector, (the previous position is designated O, because it leaves a receptor unconnected). RC is connected to the second, and so on.

On being set in the entrance of the maze on the test trial the machine is set to find the goal. The hierarchical arrangement of relays is now switched on and some receptors again pick up signals.

As has been explained, the receptors when they are stimulated send a pulse to the trolley through that portion of the hierarchy to which they are attached, but only when this portion is switched on at the same time. A portion is switched off as soon as another portion supraordinate to it has passed a pulse from the receptor, attached to it, to the trolley. Hence the trolley will always steer towards the receptor nearest to the goal and bring itself by this means within the range of stimulation of others even nearer until it reaches the goal receptor.

When the trolley has learnt two mazes with a common receptor and two separate goals the one receptor is attached to both hierarchies perhaps with a different position in each. When either goal is then selected, the switching on of the hierarchy

subordinate to that goal will be transmitted through the common receptor to portions of the other hierarchy. Only those portions, subordinate to that to which this common receptor is connected, will be switched on. Hence the trolley will find the shortest way to the goal selected in whichever entrance it is set.

The system consists of six uniselectors and twenty-three relays. The uniselectors have six wipers each, and are ordinary P.O. equipment. Only three positions are employed on each. Smaller ones would thus be more suitable, were they obtainable. The relays are similarly large P.O. relays with four "make" and one "break" contact. These also are often not all used, but the writer and constructor was led to use them because they were cheap and readily available and because he believes that he would have become intolerably confused had he used more than one type.

The system is such that it could be expanded without a loss in efficiency. For instance, even the time taken to calculate a choice or reach a decision (when there is no conflict) does not vary with the complexity of the operation (and the size of the system); it takes the machine no longer to select the path to a goal whether it has to combine information acquired in separate situations or not. This, with the number of complex properties which this relatively simple system has in common with a mammalian organism is encouraging from the point of view of theoretical psychology.

Reference

Deutsch, J. A. (1953). A new type of behaviour theory. *Brit. J. Psychol.*, **44**, 304–317.

FRUSTRATION AND STEREOTYPED BEHAVIOUR IN HUMAN SUBJECTS

BY

LEONARD C. T. JONES

From the Psychological Department, University College, London.

The investigation was designed to show the effects upon behaviour of three different durations of frustration, and two degrees of motivation during the frustrating period. Frustration was induced in 144 subjects by setting them the task of "learning" an insoluble temporal maze; they had to record their responses by pressing on one or other of two morse-type keys. Its effect was measured in terms of:

(1) The time taken to learn a soluble maze introduced, without the subjects' knowing it, by changing the system of "rewarding" responses from one based on chance to one based on the constant repetition of a pattern requiring the responses: Left—Right—Right, for its solution.
(2) The tendency for the subjects to show stereotypy of behaviour by responding on the same key for a number of trials in succession without reaching a solution.
(3) The pressure exerted on the response keys, which was taken to be a measure of vacillation.

Predictions that the time taken to learn the soluble task, and the stereotypy, would increase in direct relation to the duration of the frustrating period and the degree of motivation were tested.

It was found that, while there was an immediate increase in both the time taken to learn the soluble task and the stereotypy after a short period of frustration, a point was reached under conditions of prolonged frustration after which no further increase occurred and some adjustment to the situation was shown.

Some confirmation of the effect of increase in motivation in the direction predicted was obtained in all cases, but the differences were not statistically significant. The rankings of the subjects according to the time taken to learn the soluble task and the degree of stereotypy were found to be closely correlated. A definite tendency towards increased vacillation of response was seen in many of the records during the period when frustration might have been expected to have been at its peak.

These findings are discussed in relation to Maier's theory of frustration and to Selye's concept of a "general adaptation syndrome." The latter theory is more suitable for the interpretation of the results of the present investigation.

I

Introduction

The study of behaviour under conditions of stress and frustration has assumed considerable importance in behaviour theory, especially in relation to the problem of explaining behavioural maladaptation. In the present investigation frustration was produced in a situation in which response possibilities were limited and easily measurable. Frustration was defined as an intervening variable, the definition being taken from Underwood (1949):

> "Frustration is that which leads to the deviant behaviour which is observed as a result of the blocking of, or interference with, a goal-directed behaviour sequence."

With such a definition, an individual can only be said to be frustrated when the blocking of a drive by some obstacle leads to some form of "deviant" behaviour, or to a change in the anxiety level of the individual.

Previous investigations have shown that the behaviour resulting from frustration shows two main characteristics; firstly, it is more *stereotyped*, showing less plasticity

in a learning situation, and secondly, this stereotypy tends to persist during the period of frustration and for some time afterwards, that is to say, it shows signs of *rigidity*. But these are also characteristics of behaviour occurring as a result of any form of stress, and for this reason no stress-producing stimuli, other than the frustrating situation were employed, and in particular, the subjects' performance was in no way criticized, disparaged or ridiculed.

The insoluble problem was used as a means of producing frustration, and the performances of groups of subjects at a learning task following the frustrating situation were compared for three different durations of the insoluble problem and two degrees of motivation of the subjects during the frustrating period. Behaviour deviations were shown to occur under the conditions of stress produced by the problem, but under conditions of prolonged stress a diminished degree of stereotypy suggests that some adjustment was made to the situation.

II

Research Design and Procedure

Research Design.

The research design is presented in Table I and consisted of the following phases:

(1) A preliminary trial period of thirty seconds during which the problem was insoluble (five seconds were allowed for each response, so that six responses could be made in this first phase). Responses were "rewarded" on a random basis. The first phase gave the subjects the opportunity to become familiar with the apparatus and experimental situation, but it was essential that it should be short lest it should itself prove to be frustrating.

TABLE I

Research Design

		Phase				
Group	*N*	1	2	3		4
I	24	Trial period	Rest period	Insoluble problem: Non-incentive	5 trials	Soluble problem
II	24	"		Insoluble problem: Incentive	5 trials	"
III	24	"		Insoluble problem: Non-incentive	50 trials	"
IV	24	"		Insoluble problem: Incentive	50 trials	"
V	24	"		Insoluble problem: Non-incentive	100 trials	"
VI	24	"		Insoluble problem: Incentive	100 trials	"

Three different effects of frustration on behaviour were analysed: (i) the time taken to learn the maze following the frustrating situation (Phase 4); (ii) the degree of stereotypy of behaviour throughout phases 3 and 4 of the experiment. A succession of six responses or more to the same key was considered to be evidence of stereotypy; this criterion was chosen because under conditions of normal learning without any prior frustrating situation it was found that nearly all subjects could solve the problem after responding on one key for this number of trials; (iii) the degree of vacillation in making responses during phases 3 and 4, measured by variations in the pressure of the subjects' hand on the response keys.

(2) A rest period of two minutes during which the subjects were allowed to ask any questions about the procedure which had not been explained sufficiently in the instructions.

(3) An insoluble problem phase in which subjects of three main groups were given an insoluble temporal maze of durations 5, 50 and 100 trials, respectively, responses being "rewarded" on a random basis. Each group of subjects was sub-divided, members of each sub-group being given different instructions designed to create differing degrees of motivation.

(4) A soluble problem phase, the soluble problem being introduced, without the subject's knowing it, by changing the method of rewarding responses from a randomly determined sequence to one based on the constant repetition of a pattern requiring the responses: Left—Right—Right, for its solution. The criterion for the learning of the problem was five correct repetitions of the response pattern: Left—Right—Right; i.e. fifteen responses in all.

The following hypotheses were tested:—

(1) The length of time which it takes to solve a problem following a period of frustration increases in direct relation to the duration of the period of frustration.

(2) The length of time which it takes to solve a problem following a period of frustration increases in direct relation to the degree of motivation of the subject during the period of frustration.

(3) The degree of stereotypy of behaviour increases in direct relation to the duration of the period of frustration.

(4) The degree of stereotypy of behaviour increases in direct relation to the degree of motivation of the subject during the period of frustration.

Apparatus.

The subject was seated before a screen on which a red and a white light were displayed; fixed to a table in front of the screen was the apparatus on which his responses were made. The circuit for the red light was completed and broken every five seconds by a timer, this light acting as a signal for the subject to make a response. If a correct response was made—in accordance with the pattern which was set by the experimenter—the white light flashed on; if an incorrect response, a buzzer sounded. On the assumption that the light and buzzer gave information as to the pattern of responses to be learned, the subject attempted to eliminate those responses which resulted in the buzzer.

Responses were made by pressing down one or other of two keys which were built into a box along the top of which the subject could rest his arm comfortably, with the first and fourth fingers (forefinger and little-finger) on the two keys which protruded from the end of the apparatus. The keys were supported on a movable platform, concealed in the apparatus, which rested on a rubber bulb. Variations in the pressure exerted on the keys were conveyed through this bulb by means of a rubber tube to a tambour and pen recording on a drum-kymograph; the maximum distance through which the response keys could move was approximately two inches. The subjects rested their fingers on the keys throughout the experiment and so a continuous record of the degree of vacillation during each response was kept.

The two response keys were connected to the light and buzzer, respectively, through a mercury reversing switch by means of which the problem to be learned was set. A full record of the responses made—i.e. left or right—for each trial was kept by two pens of a four-pen tape polygraph. The third pen recorded the number of responses made by the subject for which the buzzer sounded, those responses indicating errors to the subject, while the fourth pen recorded the interval between each flash of the red light.

Subjects.

One hundred and forty-four subjects were assigned on a random basis to the six experimental groups each of which contained an equal number of men and women. The subjects were undergraduates of the University of London who volunteered to co-operate in the research and for this reason it may be assumed that they represent a fairly homogeneous group so far as intelligence level and socio-economic status were concerned. Very few of the subjects were taking courses in psychology or the social sciences. The median age of subjects was 21 years, with a range of 17 to 32 years.

Procedure.

At the outset of the experimental session one of the following two sets of instructions was given verbally to the subject, according to whether he had been assigned to a "non-incentive" or an "incentive" group.

Non-incentive groups:

"I am going to set you a problem on the apparatus in front of you. This problem will be used on other people and at present I am only concerned with finding its difficulty by trying it out on a small sample group.

"I want you to rest your arm on the box in front of you with your first and fourth fingers resting on the two keys (demonstrated by the experimenter). Your fingers should be resting on the keys the whole time. Every five seconds the red light on the screen will flash on. Each time this happens I want you to press down one or other of the two keys. Your job will be to learn a pattern which repeats itself continuously and which I am going to set you on those two keys. If you press down the correct key in accordance with this pattern, the white light in the centre of the screen will come on, if you press the wrong key then a buzzer will sound. Try to learn the pattern as quickly as possible, there is no time limit."

Incentive groups:

"This is an intelligence test which is often used in selecting college students in the United States.

"I want you to rest your arm on the box in front of you . . . (The instructions continue as above, but conclude as follows): . . . if you press the wrong key then a buzzer will sound. The time which it takes to complete this task is a good indication of intelligence; you should be able to solve it in two minutes."

At the conclusion of the experimental session, subjects of all groups were asked the following questions:—How did you go about the task of solving the problem? Did you change your method of solving the problem? Did the problem itself seem to change in any way? Did you suspect that the problem could not be solved at all? What do you consider the purpose of this experiment to be? What am I trying to find out?

III

Results

The effects of frustration on the learning time for the soluble problem.

Two subjects from the group undergoing 50 trials frustration without incentive, and three from the group undergoing 100 trials with incentive, withdrew from the experiment without having learnt the problem.

Positive skewness is evident in the distributions of the six groups, and in order to test hypotheses 1 and 2 an analysis of variance was carried out on the logarithmic transformation of the scores. This transformation was advisable since the data were of the same kind as might be obtained for a "survival" distribution and it might, therefore, be assumed that they were quasi-logarithmic in form. The results of this analysis of variance are given in Table II. Student's t tests for small samples were calculated for the differences between individual groups.

The comparisons of groups with 5-trials frustration and groups with 50- and 100-trials frustration shows an increase in the learning times for the soluble problem significant at 0·0005 level of confidence. The differences between groups with 50 trials and 100 trials are not significant and in one case are in the direction opposite to that put forward in hypothesis 1. Thus it was necessary to reject hypothesis 1 in its present form, there being no direct relation between the increase in learning time and the duration of the frustrating period.

Differences between the mean learning times of groups working under the conditions of "no incentive" and "incentive" all occurred in the direction predicted in

hypothesis 2, and the overall F ratio (Table II) shows that this effect is significant at the 0·05 level of confidence. Tests between the groups taken in pairs, however, are non-significant. It may be concluded, therefore, that the results give some support to hypothesis 2.

TABLE II

VARIANCE TABLE

Source	*Deviance*	*d.f.*	*Variance estimate*
Incentive	0·35838	1	0·35838
Frustration	5·79584	2	2·89792
Interaction	0·06125	2	0·03062
Individual differences	9·72901	138	0·07050
Totals	15·94447	144	
Interaction F = 0·0434			n.s.
Incentive F = 5·084	$n_1 = 1$	$n_2 = 138$	$P < 0·05$
Frustration F = 41·105	$n_1 = 2$	$n_2 = 138$	$P < 0·001$

The effects of frustration on the degree of stereotypy of behaviour.

While an increase in the learning time for the soluble problem may be thought of as an indirect measure of stereotypy and rigidity, a more direct method of measurement is possible. Responses of six trials or more to the same key were taken as evidence of stereotypy; in practice, such stereotypes ranged from 6 to 75 responses in duration. While this form of response was sometimes made in an attempt to solve the problem, an examination of the results and in particular of the verbal reports of subjects shows that in the majority of cases the subjects did not limit their responses in this way for rational motives and, indeed, were often unaware that they had been responding on only one key for any length of time.

Two comparisons were made between experimental groups, one between the mean duration of the stereotyped responses for each group, and the other between the proportion of stereotypes to the total number of responses made by members of each group. The second measure of stereotypy has the advantage of taking all the responses, stereotyped and non-stereotyped, into consideration, while the first only selects the stereotyped responses for comparison.

The F ratios for the effect of duration of frustration (Table III) on stereotypy are highly significant; and on the basis of these results individual t tests for the differences between groups were carried out.

TABLE III

F RATIOS FOR STEREOTYPY

A. Mean duration of stereotypes:				
Frustration	F = 21·853	$n_1 = 2$	$n_2 = 138$	$P < 0·001$
Incentive	F = 0·643			n.s.
B. Mean proportion of stereotyped responses to total number of responses.				
Frustration	F = 19·248	$n_1 = 2$	$n_2 = 138$	$P < 0·001$
Incentive	F = 1·857			n.s.

As with the learning times discussed above the differences were in all cases significant ($P < 0{\cdot}025$) between groups with 5 trials and groups with 50 or 100 trials frustration. Differences between groups with 50 and 100 trials frustration were, however, in both cases in the direction opposite to that put forward in the hypothesis. While stereotypy increases significantly with an increase in the duration of the period of frustration from 5 to 50 trials, a further increase to 100 trials does not lead to greater stereotypy, and the results suggest that under conditions of prolonged frustration the degree of stereotypy actually decreases. Hypothesis 3 cannot, therefore, be substantiated in its present form, there being no direct relation between the duration of the frustrating period and the degree of stereotypy of behaviour.

The differences between the incentive and non-incentive groups again occur in the direction predicted, but none of these differences is significant and no conclusive support can be given to hypothesis 4.

The relation between increased learning time for the soluble task and increased stereotypy of behaviour.

The increase in learning time and in the degree of stereotypy of behaviour under frustrating conditions were both used as measures of the disorganization of the subject's behaviour. In order to determine the agreement between these criteria, subjects were ranked according to their "scores" on both criteria and a correlation between these rankings was calculated using Spearman's "rho" statistic. Since "rho" in this case, is a measure of agreement between two rankings for which there is no known objective order the form "rho b" given in Kendall's (1948) treatise was used. A rank correlation value of $+ 0{\cdot}34$ was obtained between the variables of learning time and the proportion of stereotyped responses to other responses, a critical ratio of 4·0 affirming a high level of significance.

Vacillation of response.

The degree of vacillation of response was obtained through the records of the pressure made by the subjects on the response keys. The records of those subjects showing a high degree of continuous vacillation (approximately 41 per cent. of the total number of subjects) tend to reveal three phases:

(*a*) an initial steady period with little sign of vacillation;
(*b*) an intermediate period with fairly marked vacillation; and,
(*c*) a concluding period in which the amount of vacillation falls to zero.

There is, then, evidence of increasing vacillation during the frustrating situation, with a tendency for this vacillation to decrease as the subject gains mastery of the problem. It was also noticed that during periods in which stereotypes were formed, the records exhibited less vacillation, this effect demonstrating a possible "adjustive" value of the stereotype in reduction of tension as was suggested by Maier (1949) in his studies of stereotyped behaviour in the rat.

Analysis of the Verbal Reports.

The verbal reports of the subjects were found to be particularly revealing as to the existence of covert anxiety and tension; feelings of uneasiness and aggression are especially evident among those subjects undergoing the longer periods of frustration of 50 and 100 trials.

From the questions put to the subjects it would seem that the insoluble nature of the task was well disguised. None of the subjects thought the problem to be insoluble, though a few stated that at some time or other during the experimental

session they had thought that it might be too difficult for them to solve. Twenty-eight of the subjects suspected that the task had changed at some time during the experimental session, and a number thought that the change might have been made because of their inability to solve a difficult problem. In general, the method chosen to solve the problem was for the subject to test out a series of hypotheses as to what the actual pattern might be, abandoning each hypothesis as it proved false. The difficulties experienced by the subjects in the solution of the problem appear to have been, first that of finding a starting point to the pattern, and, second of deciding just how long the patterned sequence might be.

Direct questions were put to some of the subjects in an attempt to find out the reasons why they had assumed a "stereotyped" response. Replies, on the whole, were vague and unsatisfactory. Many of the subjects had no idea that they had been pressing down the same key for any length of time. Others stated that the time limit of five seconds for each response was too short for them to stop and think and this had probably led to their pressing down one key continuously. Subjects in Marquart's (1948) experiment who formed stereotypes reported that they felt compelled to continue using the same basis of choice. None of the subjects in the present experiment gave any direct hint of such compulsion, but the fact that the speed at which responses were demanded was found to be excessive by some subjects and that this was put forward as a reason for forming a stereotype may give some indirect evidence for the compulsiveness of the response. What is certain, is that for the majority of subjects the stereotype did not represent a rational attempt to solve the problem; usually the subject was unaware that he had shown any rigidity of behaviour, and when confronted with evidence of a stereotype was unable to give any explanation as to why it had occurred.

IV

Conclusions and Discussion

The results of the experiment show that an immediate disorganization of behaviour occurs as a result of frustration, this disorganization being shown by increase in the learning time of a subsequent problem and by an increase of stereotypy of behaviour both during and after a period of frustration by an insoluble problem. Under conditions of prolonged frustration, however, a point is reached after which further frustration does not lead to a further increase in the time taken to learn a subsequent problem or in the degree of stereotypy.

Differences between the groups working under conditions of "incentive" and "no incentive" in all cases show that stereotypy of behaviour is increased by the conditions of "incentive," but in two analyses out of three these differences were not statistically significant.

The assumption of the experimental method that the increase in learning time and in stereotypy were both measures of the disorganization of the behaviour of subjects was tested by means of Spearman's "rho" statistic. The value of + 0·34 was obtained when this correlation was carried out between the variables of learning time and the proportion of stereotyped responses to other responses. It is possible that this coefficient might have been considerably higher had more of the subjects shown evidence of stereotypy.

Finally, there was a definite trend towards increased vacillation over responses during that period of the task when frustration might be considered to have been at its peak, this vacillation being shown up by fluctuations in the pressure of the subject's hand on the response keys. It was also observed that vacillation did not

occur to the same degree during periods when the subject had formed a stereotype to one key, a fact which suggests that the stereotype represented some form of adjustment by the subject to the frustrating situation.

Hypotheses 1 and 3 that the learning time for the soluble task and the degree of stereotypy would increase in direct relation to the duration of frustration, were formulated largely on the basis of such experiments as those of Patrick (1934) on the disorganization of behaviour caused by an "emotional" response, and Maier (1949) and Marquart (1948) on frustration, all of whom found that behaviour under stress became increasingly more disorganized and stereotyped. The results of the present research, however, show that the relation between the duration of the period of frustration and the degree of disorganization of behaviour is not direct, and that under conditions of prolonged frustration the subject is likely to make some form of adaptation to the frustrating situation. From the study of the rat under conditions of frustration Knöpfelmacher (1953) has obtained similar results. Knöpfelmacher found that rats under conditions of high punishment during the frustrating situation showed a tendency to relinquish a stereotyped mode of behaviour more quickly, when a soluble problem was introduced, than those under low punishment. In his research the punishment consisted in preventing the animals from escaping from a water discrimination unit for periods of 8 seconds, for the low punishment groups, and 80 seconds for the high punishment groups.

Mackworth (1950) has also reported a recovery period in the efficiency of behaviour after a long period of stress. Mackworth found that subjects working under stress conditions produced by extreme temperatures or anoxia showed a gradual decline in the efficiency of their performance at certain visual and auditory tests, but after a period of between 20 and 30 minutes a recovery period was evident during which efficiency of work increased.

Maier (1949) claims that as a result of his experimental work on rats, there is good reason to suppose that "motivated" behaviour differs in *kind* from "frustration-instigated" behaviour. Working on human beings in a frustrating situation, Marquart (1948) states that the distribution of results from her experiments show definite signs of a bimodality which would be required to support Maier's hypothesis. In the present investigation no evidence to support such a distinction could be found. An examination of the distribution curves of the "scores" for learning time and degree of stereotypy showed no signs of bimodality and there were no other qualitative distinctions observable between the behaviour of the subjects under conditions of increasing frustration. Instead of considering behavioural deviations to be different in kind from "normal" motivated behaviour, it is suggested that behavioural deviations, like other reactions of the body to stress, undergo three stages analagous to those of the general adaptation syndrome put forward by Selye (1950) in the study of psychosomatic conditions.

(1) An alarm reaction, during which behavioural deviations increase progressively;
(2) A stage of resistance, in which there is some adjustment to the frustrating situation;
(3) A stage of exhaustion, in which the subject shows extreme fatigue or aggression, or withdraws from the problem.

When Cannon (1932) formulated his "emergency reaction" he showed the utility of a physiological mechanism for the explanation of behaviour under conditions of immediate stress. Selye's general adaptation syndrome is an extension of Cannon's emergency reaction to situations of prolonged stress and may well prove to be the best theoretical basis for behavioural studies of stress, frustration and fatigue.

This paper is an abridged version of a thesis—"The Effects of Frustration on Learning in Human Subjects"—submitted and approved for the Degree of Doctor of Philosophy in the Faculty of Science, University of London, 1952. The author would like to express his appreciation and gratitude to Professor R. W. Russell who supervised the research.

References

Cannon, W. B. (1932). *The Wisdom of the Body.* New York.

Jones, L. C. T. (1952). The Effects of Frustration on Learning in Human Subjects. Ph.D. Thesis, University of London Library.

Kendall, M. G. (1948). *Rank Correlation Methods.* London.

Knöpfelmacher, F. (1953). The rigidity of position stereotypes in the white rat under varying conditions of reward and punishment. Ph.D. Thesis, University of London Library.

Mackworth, N. H. (1950). *Researches on the Measurement of Human Performance.* London: H.M. Stationery Office.

Maier, N. R. F. (1949). *Frustration—the study of behaviour without a goal.* London.

Marquart, D. I. (1948). The pattern of punishment and its relation to abnormal fixation in adult human subjects. *J. gen. Psychol.*, **39**, 107–144.

Patrick, J. K. (1934). Studies in rational behaviour in emotional excitement: II, The effect of emotional excitement on rational behaviour of human subjects. *J. comp. Psychol.*, **18**, 153–195.

Selye, H. (1950). Adaptive reactions to stress. In: *Life Stress and Bodily Disease*, ed. Wolff, H. G., Proceedings of the Assocn. for research in Nervous and Mental Diseases, **29**.

Underwood, B. J. (1949). *Experimental Psychology.* New York.

CAN A TRAINED SUBJECT JUDGE HIS AUDITORY SENSITIVITY?

BY

MICHAEL WERTHEIMER

From the Psychological Laboratory, Wesleyan University, Middletown, Conn.

Auditory thresholds were measured daily on three subjects over an extended period of time, and the subjects postdicted whether the level and variability of the day's measured thresholds were higher or lower than on the preceding day, without their having knowledge of their performance. It was found that their judgments agreed with the obtained psycho-physical data better than would have been expected on the basis of chance alone. These results raise the question of the possibility that a subject's understanding of, or hypothesis concerning, the experimental outcome may influence his behaviour.

While a negative answer to the question which is the title of this note might lead an investigator in the field of sensation to breathe more easily, an answer in the affirmative would raise quite serious problems. In an attempt to answer the question experimentally, absolute auditory thresholds for clicks were measured monaurally on three subjects daily at the same time of day for 46 consecutive days. As soon as a subject's threshold had been measured, but before he saw what his day's performance had been, he tried to estimate his performance relative to his performance on the previous day. These "postdictions" were made both with respect to the value of the absolute threshold, i.e., whether it was higher or lower than on the previous day, and with respect to the variability of response, i.e., whether the standard deviation of the six samples taken daily was larger or smaller than on the previous day.

These experiments were performed in the spring of 1950 at the Harvard Psycho-Acoustic Laboratory; a detailed account of the procedure has been given elsewhere (Wertheimer, 1953). All three subjects had had extensive prior experience as subjects in auditory psycho-physical experiments.

TABLE I

Relation Between True and Estimated Levels of the Absolute Auditory Threshold for Clicks

True	*Judged*	*Subject D*	*Subject N*	*Subject M*	*Total*
H	H	12	8	16	36
H	L	4	10	5	19
L	H	4	1	8	13
L	L	13	13	12	38
Total		33	32	41	106

Tables 1 and 2 show the results of the analysis of the post-diction data. Entered in the tables are the number of times that each subject gave a judgment of "higher" or "lower" when the actual measurements indicated that the threshold or variability were higher or lower. For Table I, if the threshold was higher than on the previous day, and the subject postdicted "higher," HH is entered; if the threshold was higher, and the subject postdicted that it was lower, HL is entered; LH and LL have comparable meanings. Table II is constructed in the same way, except that "higher" and

"lower" refer to the standard deviation of the responses. On occasion the subject said, "The same as yesterday," or "I don't know;" therefore the vertical total of the table columns for each subject is less than the 45 possible.

TABLE II

Relation Between True and Estimated Levels of the Standard Deviation of Threshold Measurements

True	*Judged*	*Subject D*	*Subject N*	*Subject M*	*Total*
H	H	5	12	6	23
H	L	2	2	6	10
L	H	4	6	4	14
L	L	3	5	13	21
Total		14	25	29	68

Table I shows that 74 times out of 106, the subjects' estimates of the level of the threshold were correct, while they were incorrect only 32 times. The estimates of variability in Table II similarly were correct 44 times out of the total of 68 postdictions, incorrect only 24 times. In the case of both tables, the subjects' combined estimates of their performance differed from chance at a significance beyond the 1 per cent. level, by testing the difference from chance against the standard error of the proportion.

The subjects felt, introspectively, that they based their postdictions on many different kinds of cues, such as how loud the clicks seemed to be at and above threshold, what the perceptual quality of the clicks was, how often click quality changed during a measurement, how attentive the subject felt, and what the fluctuations in attention seemed to be.

Whatever the bases for these estimates, the fact that a subject has some ability to tell how he is doing in such an experiment raises some interesting questions, including the possibility of conscious or unconscious intent if he has some knowledge or hypothesis about the design or outcome of the experiment. It is, for example, not inconceivable that a subject's responses would be different if he expects the quantal hypotheses to hold, from what they would be if he believes the phi-gamma curve to be a more adequate description of the threshold psychometric function. In studies of cyclical changes in threshold, a subject may unwittingly be producing some regular cycles in his response sequence if he believes that cycles exist, while such cycles might not occur if the subject is convinced that threshold variations are random.

It would seem, then, that the use of a subject who is sophisticated about sensory research and theory would be inadvisable. Yet most threshold research requires a subject who is highly trained in the discriminations required for precise determinations. It is, unfortunately, rare to find a subject experienced in one without his being experienced in the other. Although it may be possible to use naive but well-trained discriminators as subjects in an occasional research design, in the overwhelming majority of contemporary studies, the investigator using trained subjects in psychophysical studies would do well to keep in mind the possibility that the subject's understanding of the experiment may not be unrelated to the results produced.

Reference

Wertheimer, M. (1953). An investigation of the "randomness" of threshold measurements. *J. exp. Psychol.*, **45**, 294–303.

DISCRIMINATION OF SIZE BY SIGHT AND TOUCH

BY

R. P. KELVIN

From the Psychological Laboratory, University College, London.

The experiment was designed to discover whether accuracy of discrimination differs when the discrimination is made in only one sense from when it is made across two senses. The evidence produced by the experiment does not warrant one to conclude that such a difference exists. The absence of such differences has been considered in relation to more general perceptual problems.

Introduction

This paper reports the results of an investigation into the accuracy with which human beings can make discriminations in size between two objects of which one is seen and the other is held in the hand. There are at least two good reasons for considering discriminations which involve more than one sense. In the first place almost any information about the sensory capacity of normal people may turn out to be a useful guide in the clinical assessment of injury and impairment. The second ground for interest is more general. There are some physical characteristics of objects which may be apprehended by more than one sense. For instance the size of an article may be perceived both by sight and by touch. Any theory about the perceptual organization and processes of either of these two senses by itself will be incomplete unless it also indicates how they can collaborate. Conversely, some helpful knowledge may be gained about each sense by studying instances of such collaboration.

The word "sight" rather than the more technical term "vision" has been used throughout this paper to emphasize that the main purpose of this study was the accuracy of discrimination in the everyday activity of looking at (divers) objects. Similarly "touch" has been used to refer to the everyday activity of perceiving things by touching them, feeling them and holding them in the hand. In this way "touch" includes both tactile and kinaesthetic processes. These definitions of "sight" and "touch" distinguish perhaps rather sharply between the performance of the subject and the physiological processes which subserve that performance. The distinction is deliberate, because there is no intention to investigate here the central or peripheral mechanisms which mediate perception with either or both senses.

Very little work has been done on discrimination across two senses. What has been done has concerned itself primarily with the P.S.E., that is, with whether things perceived by touch appear larger or smaller than things seen (Jastrow, 1889). Such work is not really relevant to the present inquiry which concerns *accuracy* of discrimination. It seems, therefore, most useful to state at once the problems of this experiment.

Aims and Basic Assumptions

The general aim of this paper was to discover whether the degree of accuracy of size discrimination across the two senses of sight and touch differed from the degree of accuracy achieved by each sense when it acted on its own.

Accuracy was held to be most usefully assessed on the basis of the Standard Deviation of two categories of judgment obtained under a Constant Method procedure.

The overall problem had to be solved by means of a number of subsidiary and more specific investigations:

(*a*) Accuracy of discrimination had to be determined for sight and touch, each acting by itself.

(*b*) Since the receptors (hand or eye) are available on both sides of the body, and since they can function either singly or in conjunction, the possible effects of laterality had to be kept in mind.

(*c*) In cross-modal discriminations, the results might be disproportionately affected either by the sense which perceives the standard or the sense which perceives the comparisons. Therefore both senses must act in both capacities.

In order to get some coherent pattern out of these second-order investigations the following assumptions were made:

(*a*) Discrimination in human beings requires the efficient working of both peripheral receptors and central nervous mechanisms.

(*b*) Each system of receptors is connected with at least one cortical centre. Cortical centres are connected with one another.

(*c*) *Accuracy* of discrimination decreases with every increase in the number of neural centres which may be assumed to subserve the discriminatory processes.

It is important to realize that these basic assumptions simply provide the framework within which the particular hypotheses of the experiment can be formulated. They are deliberately naive physiologically. They have been stated like this primarily to avoid begging questions concerning lateral or sensory dominance. Neither one side of the body, nor one sense is favoured by these assumptions. If such dominance exists, the fate of the hypotheses to which these assumptions give rise would display it. For these assumptions lead to the prediction that there will be three levels of accuracy arranged in a hierarchical order. Position in the hierarchy is determined solely by the relative simplicity or complexity of a given combination of receptors.

(1) The best performance would be predicted for the condition in which the same receptor (e.g. the right hand) perceived both the standard and the comparisons.

(2) Next in order would be the condition in which two stages were involved in the perception; i.e. where the standard and the comparisons were presented on the same side of the body, but to different senses, or where the standard and the comparisons were presented to the same sense but on opposite sides of the body.

(3) The condition for which least accuracy is predicted would be that in which the standard and the comparisons are presented to different senses and to opposite sides of the body (e.g. to the left hand and the right eye).

Two perceptual conditions are not covered by these three levels. It is not possible to predict how accuracy of discriminations involving both sides of the body simultaneously will compare with the degree of accuracy achieved by one side acting alone. However, it is possible to predict that in such bilateral discriminations, cross-modal discrimination will be cruder than discrimination within the same sense.

These three predicted levels give rise to two classes of hypotheses.

Class I: These hypotheses stipulate that there will be differences between different levels of receptor arrangements. They predict that the mean S.D. for the first level will be smaller than that for the second, and that the mean S.D. for the third level will be larger than that for the first two. Also within this class comes one prediction of *no* difference between contra-lateral comparisons in the same sense and ipselateral comparisons using both senses (these are the two types of condition that separately make up the predicted second level of accuracy).

This is the class of hypotheses which is of primary psychological interest. The results of the tests of these hypotheses can act as a basis for comparing cross-modal discrimination with the discriminatory capacity of one sense.

Class II: The hypotheses in this class are more statistically necessary than psychologically significant. They are all Null Hypotheses stipulating that there is *no* difference between the mean S.D. for any of the types of conditions which make up a given level of complexity.

In order of statistical procedure Class II hypotheses must be tested before those of Class I, since the latter can only be tested after it has been shown that each level of complexity is itself homogeneous.

It is not necessary to go into the details of the constituent hypotheses of each class. There were seven hypotheses of Class I and six of Class II. It was decided to reject a Null Hypothesis if F exceeded $P = 0{\cdot}05$. It was also decided to accept as significantly different the difference of means where t exceed $P = 0{\cdot}05$.

Design

(Throughout the remainder of this paper, the sense perceiving the standard will be written first, that perceiving the comparisons will be placed second.)

1. There are four possible combinations of receptors in terms of modalities:

(*a*) Touch—Touch (*b*) Touch—Sight
(*c*) Sight—Touch (*d*) Sight—Sight.

2. There are five possible combinations of receptors in terms of the side of the body on which those receptors may be located:

(*a*) Left—Left (*b*) Left—Right
(*c*) Right—Right (*d*) Right—Left
(*e*) Both—Both.

Each combination of modalities may be investigated with every combination of laterality. There are, therefore, $4 \times 5 = 20$ conditions to be investigated for variations in degree of accuracy.

These investigations were carried out with five subjects (2 males, 3 females) being used for each condition (cell). Each subject was tested for the condition of one particular cell only, and only one test was made with each subject. On the assumption that the subjects formed a random sample, the design permitted a straightforward application of the analysis of variance, with a minimum number of the complications that tend to confuse this procedure in psychological investigations, especially where repeated tests are made on the same subject.

Apparatus

The Stimuli:

The stimuli were a series of discs, made from aluminium 3·5 mm. thick. Their diameters ranged from 57·5 mm. to 62·5 mm., in step intervals of 0·5 mm. The size of the standard disc was 60 mm., and there were two discs of that diameter.

The weight of each disc was made to be equal to the weight of the smallest disc. This was done by scooping out a thin layer of metal from one surface of each of the larger discs. The amount of metal scooped out of any one disc was equivalent to the amount of metal by which that disc was larger than the smallest disc.

A thin strip of steel was fitted on the scooped out surface of each disc.

The Setting:

(*a*) For presentation to sight: The discs were placed on a wooden board which had been covered with green baize. This board was placed before the subject in the frontal-parallel plane, about 1 metre from the subject's eyes.

The board measured 46·5 × 31 cm. A magnet was fitted on the back of the board, 10·5 cm. from the top, and 15·5 cm. from the side of the board. Every disc could be held on on to the board by placing it so that the steel strip faced the magnet which was covered by a thin layer of baize only. Two such boards were made to enable easy presentation of the standard and the comparison stimuli in some of the visual experiments.

(*b*) For presentation to touch: The subject had to put his hand through a screen, 58·5 cm. × 33 cm. with an opening of 45 cm. × 11 cm. at its base. This opening was sufficiently wide to allow ample movement of the hands and lower arms. The height of this screen was low enough to enable the subject to have a clear view of the discs on the boards whenever this was necessary. In order to keep the general visual field as uniform as possible, this screen was also covered with green baize.

On the back of this screen there were fixtures for a light and a thin metal rod.

The light was used in the visual tasks. It illuminated the discs on the boards without casting any noticeable shadows.

The rod was fitted so that it projected above the middle of the screen. When the rod was in position it acted as the axis for a swivelling occluder. This occluder consisted of two boards, 30 × 23 cm. fixed at 90° to each other on a thin wooden block. They were covered with green baize. A hole was drilled into the block. The occluder could thus be made to fit over and pivot on the rod projecting above the screen.

By using the occluder it was possible to prevent the subject from seeing either the right or the left, or even both boards, e.g.:

(1) Suppose that the subject has to make comparisons by using his left eye only: he would be seated in such a way that the occluder obscured all views of the boards from his right eye.

(2) Suppose now that the subject had to make comparisons between his right and his left eye: he would first be shown the standard on his left, the occluder shutting off the right eye; the occluder would next be moved 45° and thus obscure all view of the boards for both eyes; it would then be swung completely over and so obscure the view for the left eye while allowing the right eye to see the disc.

By using this device, it was possible to reduce to a minimum the fatigue which comes either from the wearing of eyeshields or from the repeated opening and closing of one or the other of the two eyes. It also reduced to a minimum the optical distortions due to changes in light adaptation (Horowitz, 1949).

Procedure

Presentation of the stimuli:

The stimuli were presented in accordance with a Constant Method Procedure. The standard stimulus had a diameter of 60 mm., and there were four larger and four smaller discs for comparison. These comparison discs measured 57·5 mm., 58·5 mm., 59 mm., 59·5 mm., 60·5 mm., 61 mm., 61·5 mm., 62·5 mm. in diameter; there was also a ninth disc of 60 mm. to match the standard. Of the possible stimulus series the discs with diameters 58 mm. and 62 mm. were omitted. In most cases during pilot studies during which these discs were included, hardly any errors were made with them. Cutting these discs out considerably relieved the tedium of the experiment for the subjects. The majority of the subjects of the pilot investigation became resentful and bored by the continued re-presentation of stimuli which for the most part were obviously ("pointless") different. But as a check the penultimate rather than the extreme members of the series, 57·5 mm. and 62·5 mm., were cut out, to allow linear interpolation between the third and fifth value at each end of the series if necessary.

Each of the comparison stimuli—including the disc which matched the standard for size—was presented six times for comparison with the standard. Each of the nine comparison stimuli was presented once before any one of them was presented a second time, and so on for the second and third presentations, etc.

The order of presentation was random for the first three columns; the second three columns were mirror images of the first.

The 60 mm. disc (i.e. the standard size) was presented alternately to the receptor for the standard and to the comparing receptor.

The stimulus for the standard receptor was always presented first, for approximately 3 seconds. There was then a pause of about 6 seconds. After that the subject was given the stimulus for the comparing receptor for a period of about 3 seconds, and asked to judge this second stimulus either larger or smaller than the first. The judgments equal or doubtful were not allowed. After each judgment there was a pause of approximately 6 seconds during which the experimenter made notes of the subject's answer and selected the next two comparisons.

These times are approximations. The experimenter had a stopwatch, and in the large majority of cases the timing was adhered to very well. Occasionally, however,

TABLE I

STANDARD DEVIATION; 20 CONDITIONS, 5 SUBJECTS PER CELL

Standard Comparison	*Left Left*	*Left Right*	*Right Right*	*Right Left*	*Bilateral*
Touch Touch	1·22	1·37	0·69	1·46	1·28
	1·14	1·13	1·07	1·37	0·77
	0·98	1·18	1·23	1·45	1·84
	1·49	0·79	0·48	1·12	1·11
	1·27	1·21	1·08	1·44	1·23
Mean	1·22	1·136	0·91	1·304	1·246
Touch Sight	1·38	1·21	0·86	1·23	1·18
	1·25	1·75	1·62	1·14	0·86
	1·46	1·61	1·53	1·68	1·04
	1·56	1·79	2·10	1·48	0·90
	1·14	1·02	1·09	0·76	0·88
Mean	1·358	1·476	1·44	1·258	0·972
Sight Touch	1·34	1·44	1·14	0·85	1·31
	0·87	1·14	1·25	0·58	1·01
	1·00	1·25	1·46	1·43	0·94
	1·17	1·31	1·23	1·29	1·52
	0·85	1·17	1·07	1·43	1·11
Mean	1·126	1·262	1·23	1·116	1·178
Sight Sight	0·99	1·65	1·38	1·40	0·79
	1·02	1·35	1·55	1·27	1·19
	0·29	1·34	0·83	1·69	1·06
	0·35	1·40	1·11	1·02	1·07
	0·89	0·90	1·21	0·92	0·87
Mean	0·708	1·328	1·216	1·26	0·996

there was a slip: the stimulus disc might drop from the display-boards, the subject sometimes gave his answer quickly, at other times he would lag a little; these, and similar happenings make it impossible to vouchsave exact timing. In the main, nevertheless, the time was kept to much more closely than the experimenter had expected. Each session with any one subject lasted about 20 to 25 minutes.

Subjects:

Forty men and 60 women acted as subjects in this experiment. The subjects were volunteers and were told before volunteering the general aims of the experiment. They were none of them acquainted with the aims, methods and complications of experimental psychology. Their performance was thus not biased by preconceived notions about what happens in psycho-physical experiments. The men were mainly office staff at every level of competence, though there were some research scientists among them as well. The women were predominantly clerical workers and secretaries, but there were also some technicians and research workers among them. The subjects were taken singly, and the instructions varied with the twenty conditions under any one of which a given subject might be tested. However, all instructions stressed that the subject should attend as carefully as possible to his first impression of the stimulus; he was warned that trying to "work out" and thinking hard about the size of a disc would only make it more difficult for him to come to a decision; he was assured that he would do best by attending to first impressions.

The majority of subjects were right-handed. Those who reported that they were left handed were used for bilateral tasks only.

Results

Table I shows the S.D. for each subject, and the mean S.D. for each cell. The first two scores in each cell are the scores for the men in that cell, the remaining three values are the S.D.s of the women. Each value was computed by the Spearman summation method (*cf.* Woodworth, 1950, p. 403).

TABLE II

	Sum of Squares	*d.f.*	*Mean Square*	*V.R.1.*	*V.R.2*
Sides	6051·9	4	1512·975	1·909	1·979
Senses	5207·8	3	1735·9	2·191	2·271
Sex	0·167	1	0·167	0·211	0·218
Sides × Sex	2169·683	4	524·421	0·662	0·686
Senses × Sex	5954·7	3	1984·9	2·505	2·597
Sides × Senses	20436·1	12	1703·008	2·149*	2·23*
Sides × Senses × Sex ..	7497·32	12	624·78	0·789	0·764
Error	47535·33	60	792·25		
Total	94853	99			

Where the value in this table or subsequently is marked thus *, P = <0·05 and where ** P = <0·01.

Table II shows the three-way analysis of variance for this data. Since the triple interaction was not significant, its sum of squares was combined with that for the error term, and the resulting new mean square was used to obtain V.R.2 with 72 degrees of freedom. As the three-way analysis shows significant interaction between sides and senses, it was reasonable to proceed to test the various individual hypotheses to see whether there was a coherent pattern of differences of accuracy. It is only necessary to summarize the outcome of these tests.

All the hypotheses of Class II (i.e. all those stipulating *no* differences between the means of certain groupings of cells) were acceptable apart from one. The unacceptable hypothesis predicted that there would be no difference between any of the perceptual conditions in which the standard and the comparisons were presented to the same receptor (hand or eye). In other words the Null Hypothesis underlying that condition for which greatest accuracy had been predicted, had to be rejected. Consequently those hypotheses of Class I which were designed to test the relative degrees of accuracy of other perceptual conditions with respect to this highest level, could not be tested. However, when the remaining hypotheses of Class I were tested, none showed the predicted differences—or indeed any significant differences. The only hypothesis of this class which was acceptable was also the only one in this class which predicted that there would be no difference between the means of the two types of condition which make up the second level of accuracy (on the basis of the underlying assumptions of the experiment).

As a result of this failure to find any of the predicted differences between the various groups of cells tested, the cells were regrouped and tested for effects of lateral or sensory dominance.

Lateral Dominance: All the hypotheses were null hypotheses. All were acceptable, with one exception. The hypothesis stipulating that there would be no difference within the conditions in which the standard was presented to the left receptor, and where the comparisons were made in the same sense, had to be rejected. This meant that no test could be made between this group of cells, and the corresponding groups on the right side of the body. It was however plain that laterality did not affect crossmodal performance.

Sensory Dominance: The aim here was to test whether cross-modal performance was cruder than ipsemodal performance (i.e. performance in which the standard and the comparisons were perceived by the same modality). Each of the four conditions in terms of modalities was taken as a whole and tested for homogeneity. All but that in which both the standard and the comparisons were presented to sight, showed homogeneity. The sight-sight group had an $F = 4{\cdot}02$ for 4 and 20 d.f., which is significant beyond $P = 0{\cdot}01$. It was therefore not reasonable to proceed to test the difference between the ipsemodal and the crossmodal tasks. As a matter of interest the difference between the means of the two cross-modal tasks was tested. $t = 1{\cdot}409$ (d.f. = 48) which is not significant.

Both the Null Hypotheses based on the original assumptions, and the *post-hoc* hypotheses concerning lateral and sensory dominance thus fail to show systematic differences in the degree of accuracy achieved under the various conditions of the experiment. At the same time some hypotheses could not be tested because one of the groups between which differences were to be tested showed itself to be internally significantly inconsistent. On examining the raw scores, these internally inconsistent groups always contained the bottom left-hand cell of Table I: the condition in which the left eye perceived both the standard and the comparisons. Inspection of the values of this cell will show that it contains particularly low values for the standard deviation. This conglomeration of small S.D.s in this one cell is almost certainly responsible for the heterogeneity found within the groupings of cells in which this cell participated.

In the three-way analysis of variance, however, there is a noticeably small sum of squares for sex. So much is this the case that it was thought justifiable to examine the results of the men and the women separately in order to discover whether the results for the whole data have been obscured by a counterbalancing of the results of men and women respectively. Therefore all the hypotheses which had been formulated for the combined data were now tested separately for men and women. Very little more positive information was gained.

Men: The two-way analysis of variance showed no significant effects. The highest V.R. was for Sides (2·067, d.f. 4 and 20) and even that was not significant.

All the hypotheses of Class II were acceptable, and therefore all the hypotheses of Class I could be tested. Two of these were acceptable. One was the only null hypothesis in this class. The other showed that performance in the same sense but across opposite sides of the body was less accurate than performance in which the same receptor perceived both standard and comparisons.

There were no significant differences due either to lateral or sensory dominance. Again all hypotheses could be, and were, tested.

Women: The two-way analysis of variance showed a significant effect with respect to the Senses. (V.R. = 3·423,* d.f. = 3 and 40.)

Again all hypotheses of Class II were acceptable and all the hypotheses of Class I

could therefore be tested. Three of these were acceptable. One was the null hypothesis of this class. Of the other two

(*a*) Cross-modal accuracy was found to be significantly cruder than ipsemodal accuracy, when the comparisons were presented on the same side of the body.

(*b*) Cross-modal tasks involving opposite sides of the body were found to be significantly less accurate than ipsemodal task performed ipselaterally.

There was no indication of lateral dominance for women.

There was, however, a significant difference between all ipsemodal means and all cross-modal means.

Ipsemodal Mean : 1·084
Cross-modal Mean : 1·289
} $t = 2{\cdot}54^{**}$ d.f. = 58

It is, therefore, clear that a slightly more articulated picture emerges when the results for men and women are separated. There is, for instance some suggestion that the performance of men is more affected by the sides of the body involved in the discrimination, whereas the performance of women seems to be more affected by whether one or two senses have to be used. However, though the results for the separated data may seem slightly more positive, this does not mean that it is either justifiable, or even useful to treat the results of men and women separately. It should be pointed out that the age range of the men and the women differed significantly, but that there was no significant correlation between age and S.D.

Discussion

The first problem which arises is to decide whether it is more reasonable to interpret the results for men and women combined, or whether to follow up the results of the two sexes separately. The results for the combined and separated data are sufficiently divergent for it to be impossible to accept both. Thus some hypotheses which had to be rejected for the combined results were acceptable for the separated data. The writer has come to the conclusion that on the basis of the information available to him there is not sufficient cause to treat the results for men and women separately.

From a purely statistical standpoint there were originally thirteen hypotheses. Six of these postulated that there would be differences between the conditions tested under them. Seven of them (including one from Class I) predicted that there would be no differences among the conditions tested under them. For the combined results six out of the seven hypotheses predicting homogeneity could be accepted; only one had to be rejected. Three only of the hypotheses which stipulated differences could be tested; not one was acceptable. The combined results thus certainly show no strong tendency for accuracy of discrimination to vary systematically with the conditions under which the discrimination has to be made. This same conclusion follows from the tests of the *post-hoc* hypotheses concerning lateral and sensory dominance.

The position is similar for the separated data. Because the numbers per cell were reduced, the hypotheses covering bilateral discrimination were telescoped (i.e. all the four possible bilateral conditions were taken together and tested for homogeneity, and were found homogeneous). There were, therefore, only six hypotheses postulating homogeneity and only five predicting differences. Both for men and for women all the hypotheses about homogeneity were acceptable. For men only one, for women only two of the hypotheses concerning differences could be accepted; and it should be noted that all these hypotheses were testable. Moreover, of this group of hypotheses, those acceptable for men and those acceptable for

women were different. While at first glance this fact would support the policy of keeping the results for men and women apart, the writer believes that such a policy would be mistaken, or at least misleading.

An overall examination of the results, both combined and separated, emphasises the absence of differences rather than their systematic distribution. In the case of the separated data a total of twenty-two hypotheses was tested only three of which showed differences to exist. In view of the small number of cases per cell, and the strong indication of homogeneity over the whole table, it seems reasonable to suggest that these differences were only accidentally significant. Also, since the left eye failed to predominate in the separated data, its effect on the combined results was probably due to an accidental grouping of small S.D.s in one cell rather than to a systematic difference made more apparent by a larger number of cases.

In other words the writer feels that there is not sufficient evidence to suppose that the twenty conditions examined in this experiment show any coherent pattern of differences. If the various hypotheses had been acceptable, or if at least some of them had shown some ordered differences in accuracy between various conditions, then the small number of cases tested under each condition would have been compensated for by the consistency of the results. But one cannot make generalizations on the basis of five subjects. At the most one can state that in this experiment one or two conditions did not produce results similar to the rest. Such a statement would be an *ad hoc* comment on the experiment, with no consequences. It would in fact be tantamount to the admission that those differences might be due to sampling, and that would bring one straight back to the position the writer has chosen to adopt. However, the conditions for which differences have been found must be investigated further, so that if the differences are confirmed, they can be put forward as generalizations rather than as asides.

These results and conclusions are not really very surprising. Certainly tradition tends to favour the prejudices concerning handedness; that most persons are either right-handed or left-handed. Yet both observational and experimental investigations show that there is no simple form of "handedness." Whether in terms of the hand used in everyday activities, or whether in terms of special laboratory tasks, authors have been repeatedly forced to the conclusion that handedness must be defined in relation to a particular situation. (Cuff, 1931; Downey, 1933.)

The evidence for eye-preference is no stronger. Buxton and Crossland (1937) report that eyedness tends to vary with the test which is used to establish it. Moreover, since the eye transmits to both hemispheres it would in any case be unwise to generalize from eye-preference to the dominance of one hemisphere. There is also some evidence that the preferred eye is not necessarily the one with greatest acuity (Gahagan, 1933).

The absence of lateral dominance in the present experiment is therefore not surprising. This absence of evidence for lateral dominance only partly explains the absence of difference between ipselateral and contralateral comparisons. Though it may be permissible to infer that either side of the body is equally sensitive, one cannot therefore assume that there will be no difference between ipselateral and contralateral comparisons; it might, thus, be quite reasonable to assume that the two hemispheres and receptor mechanisms can act equally well, but that they have only rudimentary connections. However, anatomical connections are known to exist. Further, the recovery of functions after brain-lesions, both in man and animals, and the many forms of bilateral transfer, make it clear that in many discriminatory processes both hemispheres are at least partly involved; even where at first sight only one hemisphere would need to be active.

The issues raised by crossmodal discrimination are more complicated. The basic assumptions led to the prediction that crossmodal accuracy would be cruder than ipsemodal accuracy. The results for women showed that this prediction may yet be found tenable, although it had to be rejected for the combined data. But whether cross-modal discrimination turns out to be cruder than ipsemodal discrimination or not, it is quite obvious that human beings can estimate differences of size between a visually and a tactually presented object to a remarkable degree of accuracy. Speaking loosely, on the basis of the present experiment it seems reasonable to suggest that a difference of 1·5 mm. in diameter for objects varying around 60 mm. in diameter, would be detected quite easily. It has, however, almost been a maxim of experimental psychology that estimation of size is dependent on a host of cues which are ostensibly unrelated to the stimulus being judged. This belief is well illustrated by Graham (1950). Writing about experiments concerning visual perception of size, he remarks:

> "Size matches are dependent on the presence or absence of stimuli other than those which may be considered to be *the* discriminative stimuli."

These "other stimuli" are generally thought to be stimuli derived from the environment in which "the stimuli" are presented. But in view of the degree of accuracy of the cross-modal matches of the present experiment, Graham's conclusions become a little doubtful. For while it is, prima facie, reasonable to admit the influence of the background on *visually* perceived size, there is no such background to the *tactually* perceived stimulus. It is therefore difficult to explain how the combination of background and stimulus in the case of sight, produces a visual impression of size which can be matched so accurately with a tactual impression of size, which is not affected by such an extraneous factor in the perceptual situation as the concomitant background. Though in the present experiment the writer took every precaution not to introduce disturbing elements into his visual presentations, he finds it hard to believe that he eliminated all the effects that are usually attributed to the background. He finds it even harder to believe that by some fortunate accident he presented his visual stimuli against a background, and at a distance, which yielded a visual impression of size that almost precisely matched the tactual impression of size in which such a background was lacking. Instead he ventures to suggest that his results would not have been very different for a considerable range of distances and with a fair amount of variation in the area of the display-boards. This may have to be investigated experimentally, but the writer feels sure that over a middle range of distances, and in the absence of distortions deliberately introduced into the background, the results of cross-modal matches would not differ significantly from the ones which he obtained. There are two consequences to this conclusion.

In the first place, it may be worth considering cross-modal matches as a special case of constancy phenomena. For a match across two senses demands that the stimulus patterns of the two modalities have some common features; and these common features can only arise from the characteristics of the stimulus object. In view of the absence of possibly similar effects of the background on sight and touch, these common features can only be certain physical characteristics of the stimulus. The perceptual situation would therefore seem to constitute an instance of "phenomenal regression" to the "real" object.

The second consequence is concomitant with the first. It is not sufficient to maintain that "constancy" effects themselves are due to cues derived from the experimental setting. Though environmental factors may play an important part in some visual constancies, they cannot account for the form of constancy found in

cross-modal comparisons. Constancy within one modality is determined by the relative consistency of, e.g. judgment of size, for conditions which are considerably different in terms of the physical effects of the stimulus on the organism. This consistency is mostly attributed to cues from the environment. These are thought to supplement those of the primary stimulus in such a way that they compensate the changing effects of that primary stimulus due to changes in, e.g. distance. In terms of Gestalt Psychology one might express this by saying that the field, composed of the stimulus and its background, is structured so that changes in one of its aspects tend to be counteracted by simultaneous changes in the remainder of the field. But while this balancing act of the perceptual mechanism may account for the consistency of judgments in one modality, it is difficult to see how it can account for the agreement in judgment of size between sight and touch. There is no common physical field or background to the visual and tactual stimuli used in this experiment which is analogous to the field or the rôle of the background usually postulated for results showing constancy. The problem is therefore to explain how the organism achieves such a high degree of accuracy in comparing across two senses. For such matches must be considered to depend on the physical characteristics of the stimuli judged and not on the structure of the background against which they appear.

The fundamental point of interest is thus not so much how favourably ipsemodal accuracy compares with cross-modal accuracy, but that cross-modal judgments can be made at all and relatively so accurately. The writer does not begin to know what the answer is. Only certain considerations suggest themselves.

Performances involving two modalities are not an experimental artefact but an everyday occurrence. There is virtually no activity in which there is "feedback" where this feedback does not come by means of a sense other than that used in the performance. This is true of mirror-drawing, playing cricket and even talking. To ask a person to make a discrimination across two senses is not therefore asking anything superhuman. The neural mechanisms may scarcely be known, but it must be assumed that they are so constituted that the specific character of each sense does not dominate over whatever it may be that, neurologically, the senses have in common.

It is quite clear that certain characteristics of objects may be apprehended by more than one sense. This is particularly the case with sight and touch, where the same feature of the stimulus can be perceived by both senses. But each sense apprehends this feature in its own way. The writer wishes to suggest that an amended concept of the "Schema" as propounded by Head (1920) may be helpful in thinking about cross-modal phenomena. On the basis of his clinical observations Head found himself forced to conclude that there is no fixed image of the body which could act as a standard for the appreciation of postural change. Posture was continually changing in some way or other. Normal human beings nevertheless had the capacity to indicate, for example, where a certain movement had deposited their arm with respect to the position of that arm before the movement; and so with a second or third movement. Therefore, if there were an "image" which acted as a standard it had to be one which was constantly modified by the continual postural changes. Since such an entity was scarcely an "image" in the usual almost pictorial sense of that word, Head labelled this complex standard "Schema."

Similarly, appreciation of differences between two comparable stimuli presented to two senses, cannot depend on some matching of static images against each other. Instead the writer postulates that as a result of its perceptual experiences the organism develops the capacity to react in the same way to those features of perceptual situations which may be mediated by more than one sense; and that eventually it becomes able to react to those features independently of the modality which at a given moment

happens to mediate them. That is, the organism learns to react to size, for instance, but the particular sense which subserves the perception of size at a given instant may be relatively unimportant.

This concept of Schema is in no sense the concept of an image; any of the normal uses of the word "image" imply specific sensory processes. But the whole point of introducing this amended version of the "Schema" was to obviate the need to think in terms of one or other of the senses when one considers situations where several senses co-operate. In this interpretation the Schema is a manner of functioning of the organism; it is a skill which the organism develops and which enables it to make the greatest use of its sensory equipment. The physiological processes of such a Schema are a mystery to the writer. But the full implications of the concept can be established experimentally by a host of possible investigations involving different aspects of performances in which two senses—or more—take part.

This paper has been adapted and abridged from a thesis of the same title submitted for the degree of Ph.D. at the University of London. The writer is particularly indebted to Mr. J. W. Whitfield for his constant help in supervising this experiment.

This research was assisted by a grant from the Central Research Fund of the University of London.

References

Buxton, C. E., and Crossland, H. R. (1937). The concept of "Eye-Preference." *Amer. J. Psychol.*, **49**, 458–461.

Cuff, N. B. (1931). A study of eyedness and handedness. *J. exp. Psychol.*, **14**, 164–175.

Downey, J. E. (1933). Laterality of function. *Psych. Bull.*, **30**, 109–142.

Gahagan, L. (1933). Visual dominance-acuity relationships. *J. gen. Psychol.*, **9**, 455–459.

Graham, C. H. (1951) in Stevens, S. S., *Handbook of Experimental Psychology*, Chap. XXIII, New York.

Head, H. (1920). *Studies in Neurology*, Vol. II, Oxford.

Horowitz, M. (1949). An analysis of the superiority of binocular over monocular visual acuity. *J. exp. Psychol.*, **39**, 581–596.

Jastrow, J. (1889). Perception of space by disparate senses. *Mind*, **11**, 539–554.

Woodworth, R. S. (1950). *Experimental Psychology*, London.

FIGURAL AFTER-EFFECTS, RETINAL SIZE, AND APPARENT SIZE

BY

N. S. SUTHERLAND

From the Institute of Experimental Psychology, University of Oxford.

The results of earlier experiments on the question of whether figural after-effects are affected by *apparent* as opposed to *retinal* size are shown to be inconclusive. A new hypothesis is proposed namely that both factors may be responsible for producing figural after-effects, and four experiments have been made to test it. Situations were used in which the apparent sizes of the figures were determined by the size-constancy effect. It was found that where retinal sizes of test and inspection figures are the same and apparent sizes are different, figural after-effects in the direction which would be predicted on the basis of apparent sizes are obtained. It was further shown that where retinal and apparent sizes are in conflict, whether a figural after-effect is seen or not, and the direction of the figural after-effect, depends upon the balance between these two factors.

I

Introduction

A number of experiments have been made in recent years to discover whether visual figural after-effects depend upon retinal size or apparent size. Previous experimenters seem to have assumed that the two alternatives are mutually exclusive, and they have tended to use experimental designs in which the kind of figural after-effect obtained after looking at the inspection figure would differ according to whether it was being determined by the apparent or retinal size of the figures. We wish to argue that previous experiments have been inconclusive because they have started from the assumption that figural after-effects must be either wholly dependent on retinal size or wholly dependent on apparent size, and we have performed experiments to show that this is a false dilemma and that in fact both factors are operative.

In what follows we shall use the expression *inspection figure* to refer to the figure inspection of which determines the figural after-effect. In producing a figural after-effect, a fixation point is exposed with the inspection figure and the subject fixates this point for a period of about a minute. The term *test figure* refers to the figure which is exposed to the subject after the period of fixation of the inspection figure. Its contours are made to fall near the position on the retina upon which the contours of the inspection figure fell by getting the subject to fixate a fixation point exposed with it. It is the test figure which undergoes a distortion when there is a figural after-effect. For the purposes of detecting such a distortion a third figure is often exposed at the same time as the test figure but on the other side of the fixation mark, and the subject is asked either to pass some judgment on the apparent difference between this figure and the test figure, or to adjust this figure to the same size, or position as the test figure. This is the *comparison figure*. We make the assumption throughout that in figural after-effects the contours of the test figure tend to be seen as displaced away from the position occupied by the contours of the inspection figure (*v.* Kohler and Wallach 1944). The issue at stake is how far is such displacement of the test figure governed by the retinal positions of the contours of test and inspection figures, how far by the apparent positions and sizes of the test and inspection figures.

We now proceed to a detailed consideration of previous experiments on figural after-effects in relation to apparent and retinal size. Two different methods have been employed by previous experiments on this problem:

(1) Using a method first employed by Thouless (unpublished) Prentice (1950) and Hochberg and Bitterman (1951) attempted to discover whether apparent size or retinal size was determining the figural after-effect by placing the inspection and test figures at different distances from the subject. The figures used in both experiments were circles and in both cases the more distant figure was considerably larger than the nearer. (*v.* Fig. 1 and accompanying text.) In Prentice's experiment the sizes of

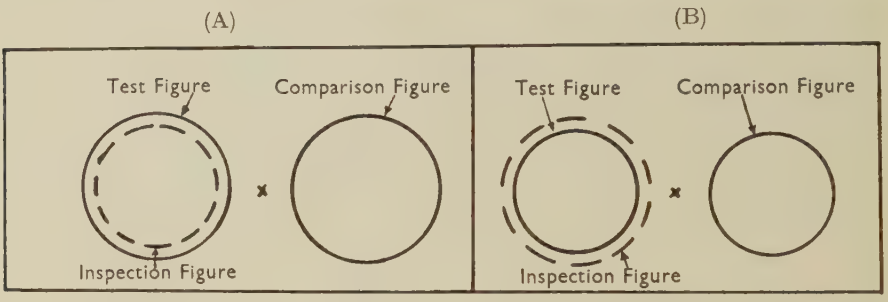


Fig. 1. Illustration of experimental design used by Prentice (1950), Hochberg and Bitterman (1951) and self (Experiment 3).

Solid outline circles represent the test and comparison figures. Broken outline circles represent the inspection figures.

In Prentice's experiment (A) depicts the relationship between *retinal sizes*—the inspection figure is retinally smaller because it is further away from the subject. (B) depicts the relationship between the *apparent sizes* of the figures—the inspection figure having larger physical dimensions than the test and comparison figures is of a larger apparent size despite its greater distance. If retinal size is the determining factor in figural after-effects then the left hand figure (test figure) should be seen as larger than the right hand figure (comparison figure). If apparent size is the determining factor then the right hand test figure should look larger than the left.

In Hochberg and Bitterman's experiment and in own experiment (3) this situation is reversed. (B) now represents relative *retinal sizes*, (A) represents relative *apparent sizes*. If retinal size determines figural after-effect the right test circle should be seen as larger than the left; if apparent size determines figural after-effect, the left test circle should be seen as larger than the right.

the figures and their distance from the subject were so arranged that while the *retinal projection* of the inspection figure fell within that of the test figure, owing to the effects of constancy the inspection figure had a larger *apparent size* than the test figure, i.e. the inspection figure was larger and further away than the test figure. In Hochberg and Bitterman's experiment this situation was reversed: the inspection figure was nearer and smaller than the test figure, and their relative sizes were such that the retinal projection of the inspection figure lay outside that of the test figure, whereas the apparent size of the inspection figure was smaller than that of the test figure. Thus in Prentice's experiment the test figure should be seen as smaller if figural after-effects depend on apparent size, as larger if they depend on retinal size, while in Hochberg and Bitterman's experiment the reverse is true. In both experiments results were obtained which unequivocally pointed to retinal size as the determining factor.

If, however, we allow that both factors may contribute to figural after-effects, in a situation in which the two factors are in conflict the result will depend presumably on

(*a*) the absolute extent to which retinal size and apparent size can contribute to figural after-effects in any situation, and (*b*) the differences in the apparent sizes and retinal sizes of the inspection and test figures in a given situation. Thus:

(*a*) It is possible that even if both retinal size and apparent size are important in determining the figural after-effect, the retinal size is the more important factor of the two and in the experiments under consideration where the two factors are in conflict any effect due to the apparent sizes of the test and inspection figures might be concealed.

(*b*) As far as can be judged, in both experiments the separation of the figures according to retinal size was about the optimal for producing figural after-effects. Since measures of constancy for individual subjects are not quoted in either of the reports of the experiments, it is impossible to say accurately what the relative apparent sizes of test and inspection figures were. Kohler and Wallach (1944) have shown that if the contours of test and inspection figure are either too close together or too widely separated figural after-effects diminish, and this may have been the case with the apparent sizes in the experiments under review. In addition to this, constancy may be diminished by prolonged fixation (though this cannot be overcome by any experimental design in which the differences in apparent size between test and inspection figures are the result of size constancy).

(2) In an earlier experiment of different design, Prentice (1947) obtained results which he interpreted at the time as favouring the hypothesis that apparent size and not retinal size was the determining factor in producing figural after-effects. His experimental design is shown in Figure 2 and the text thereto. In this experiment the inspection figure was exposed at 2 metres from the subject and the test figure was exposed for some trials at 2 metres and for others at 6 metres. If retinal size of the

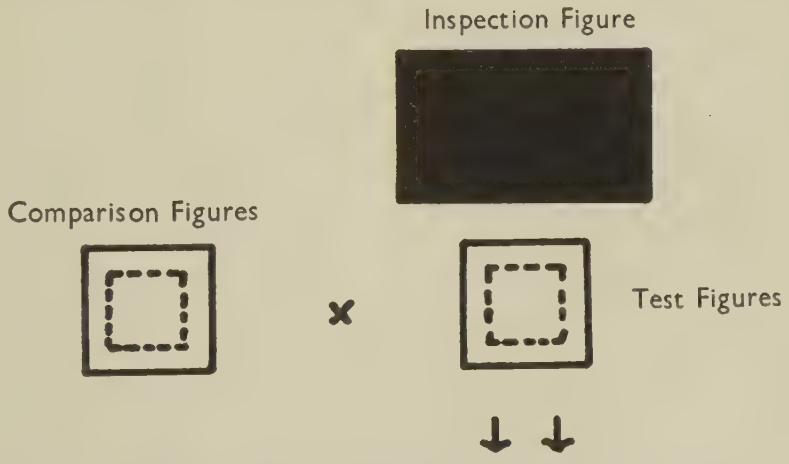


FIG. 2. Illustration of the experiment performed by Prentice (1947).

The filled in rectangle is the inspection figure. The squares with solid outlines represent retinal sizes of test and comparison figures exposed at the same distance as the inspection figure (condition 1). The squares with broken outlines represent retinal sizes of test and comparison figures when exposed further away than the inspection figure (condition 2). Arrows represent direction of displacement of test figure in figural after-effect. The issue is whether the test figure will appear displaced through the same angular distance in condition 2 as in condition 1 (retinal size operative), or through the same spatial distance i.e. a smaller angular distance (apparent size operative).

figure is the determining factor, the right-hand test square should be displaced approximately three times as much in relation to the left at a distance of 6 metres as at a distance of 2 metres. He found that in fact there was no difference in the PSE of the test and comparison squares under the two conditions. He points out, however, that this result might possibly be explained on the grounds that when the test squares are at 6 metres, the contours of the right-hand test square will, in fact, be three times as far away from the contour of the inspection rectangle as at 2 metres measured in terms of retinal size. This difference in angular separation of the test and inspection figures might lead to a decrease in figural after-effect (measured in terms of retinal displacement), and since the amount of displacement found at a distance of 6 metres was, in fact, one-third of that at a distance of 2 metres (i.e. the actual displacement of the test figure was the same in both cases), there is an alternative explanation for this result. Prentice tried to rule out this explanation by repeating the experiment with the inspection figure at a distance of 3 metres and the test figure at a distance of 3 metres, 5 metres, and 7 metres. Under these three conditions Prentice again found that the PSE for the setting of test and comparison squares remained the same. Since it seems unlikely that in all the changes of distance of the test figures made, the resulting differences in angular separation of the figures should produce exactly the same result, namely that the actual displacement of the test figure should remain the same while the angular displacement changes through varying degrees, it seems that this experiment affords some evidence for apparent size being a factor in figural after-effects. Prentice himself adopted the alternative explanation of this experiment after he had performed his 1950 experiment (described above) despite the fact that he had already attempted to rule out that explanation in the design of the experiment we are discussing. If now we treat the 1947 experiment as affording evidence for the influence of apparent size on figural after-effect and the 1950 experiment as affording evidence for the influence of retinal size, then we have no need for the alternative explanation of the 1947 experiment which Prentice himself admits to be unsatisfactory.

One important difference between the 1947 experiment and the 1950 and 1951 experiments should be noticed. In the 1947 experiment both retinal and apparent size would be working to produce a displacement in the same direction, whereas in the later experiments, in which no evidence for the importance of apparent size in figural after-effects was discovered, any effect due to apparent size would be in conflict with any effect due to retinal size.

Thus the whole issue of how far retinal size and how far apparent size influence figural after-effects can be reopened, and in reopening it we must be careful not to start from the premiss that figural after-effects are due entirely to retinal size or entirely to apparent size. Both factors may contribute, and the resulting after-effect may be due to the balance between them. If in fact perception is a two-stage or multi-stage process, then it seems likely that both factors would be of importance in determining figural after-effect, since whatever the mechanism of the figural after-effect it may be at work at each stage of the perceptual process, and the fact that figural after-effects have been shown to occur in the third dimension (Kohler & Emery 1947) indicates that figural after-effects do in fact occur at some late stage in the perceptual process.

The issue is one of considerable importance not only for arriving at the mechanism behind figural after-effects but for arriving at the brain mechanism responsible for apparent size and in particular size constancy. Thus if Kohler's explanation of figural after-effects in terms of figural currents is correct, then given his theory of apparent size, he must predict that the apparent size of figures will have some influence

on figural after-effects since the mechanism he gives for apparent size involves brain fields and a topographical projection, and within such a mechanism satiation will presumably occur. On the other hand if Osgood and Heyer's theory of figural after-effects is correct, then the discovery that figural after-effects are in part determined by apparent size will limit the kind of mechanism we can postulate to explain apparent size and the constancies. The present series of experiments was undertaken in an attempt to discover whether apparent size influences figural after-effects under optimal conditions, i.e. where no contrary effect due to retinal size would be predicted. The basic design of the experiment follows that employed by Thouless (unpublished) and Prentice (1950), but it differs from theirs in that in the crucial experiments the retinal sizes of the figures were made equal.

II

Experiments

Experiment 1. In experiments (1) and (2) an attempt was made to eliminate the influence of retinal size as much as possible by using a situation in which retinal sizes were equal, but apparent sizes were not. By this method it was hoped to isolate any effect that might be due to apparent size. In the first experiment the inspection figure was an outline circle of diameter 4 in., drawn in Indian ink, on a large white card. The test figure was an outline circle of diameter 10 in. and was exposed with a comparison figure of identical dimensions on the other side of the fixation point. Test and comparison figures were fixed to a white wall. The inspection figure was at a distance of 4 ft. 9·6in. from the subjects' eyes: the test figure at a distance of 12 ft. At these distances the angle subtended by the test and inspection figures on the retina should be the same (*v.* Fig. 3). Subjects were asked to fixate the inspection figure for a period of 45 seconds, whereupon the inspection figure was withdrawn suddenly, and subjects fixated the fixation point on the wall and were asked to say which of the two circles on the wall looked bigger to them.

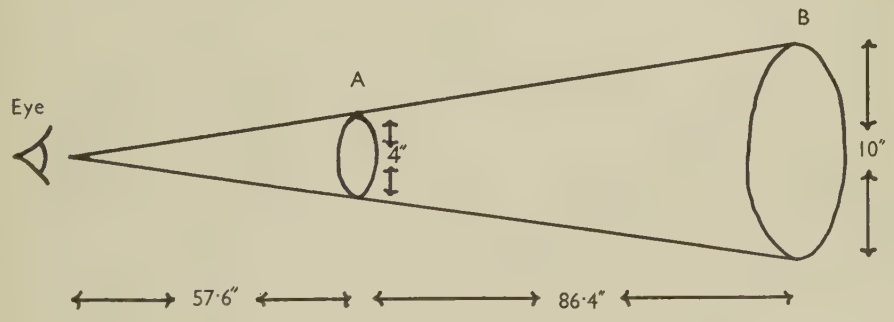


Fig. 3. The design of own experiments (1) and (2).

In experiment (1) A was the inspection figure, B the test figure. If there is a figural after-effect due to apparent size, the test figure should be seen as larger than the comparison figure, since the apparent size of A is less than B, i.e. where the inspection figure (A) is left of the fixation point the left hand of the two identical circles at (B) should be judged the bigger, where the inspection figure (A) is right of the fixation point, the right hand of the two circles at (B) should be judged the bigger. In experiment (2) B is the inspection figure, A the test figure. If there is a figural after-effect due to apparent size, the test figure at (A) should be seen as smaller than the comparison figure at (A), since the test figure (A) now has a smaller apparent size than the inspection figure (B), i.e. if the inspection figure (B) is to the left of the fixation point the right hand of the two identical circles at A should be judged bigger, if the inspection figure is to the right of the fixation point the left hand of the two circles at A should be judged bigger.

All instructions were given before the experiment began. With half the subjects the inspection figure was to the left of the fixation point, with the other half to the right. The distance between the fixation point and the edge of the circle was 0·75 in. on the inspection figure and 2·25 in. on the test figure. Each subject gave only one judgment.

Experiment 2. In the second experiment the inspection figure was the far figure, and the test figure was the near one, i.e. one circle only was first exposed at a distance of 12 ft. to right or left of the fixation point, and two figures (test and comparison) were then shown at 4 ft. 9·6 in. Subjects were asked to close their eyes for a period of just over a second on the termination of the 45 second inspection period in order to prevent them seeing the test figures before they were in position. Otherwise the procedure was exactly the same as for experiment 1.

Experiment 3. The third experiment was undertaken in order to confirm the results of Prentice (1950) and Hochberg and Bitterman (1951) i.e., in order to discover what happens when retinal size and apparent size are in conflict. In this experiment the inspection figure was the near figure as in experiment 1, but it was exposed at a distance of 4 ft. (*v.* Fig. 1).

Experiment 4. In the fourth experiment an attempt was made to determine for individual subjects the distance from the eye at which the test figures appeared equal after inspection of the figure at 12 ft. Each subject made 70 judgments on the relative apparent sizes of the two circles exposed on the test card. The distance of the test card from the subjects' eyes was varied, 7 different distances being used: 10 presentations were made at each distance in a random order. The order was varied from subject to subject. The procedure was as follows: after a preliminary inspection period of 45 seconds, the subject closed his eyes for a period of just over a second while the test figure was placed in position. On the command "Look" given by the experimenter, subjects opened their eyes fixated the fixation point on the test card and gave their judgment "Left bigger," "Right bigger," or "Equal" as quickly as possible. As soon as the judgment had been made, the test card was removed, subjects fixated the inspection figure for a period of 10 seconds and the procedure was repeated. Unfortunately, all judgments had to be made in one session. After successive groups of 14 judgments, subjects were given 10 minute rest periods. Two subjects had to be discarded because their eyes grew tired with the successive fixations before the experiment had been concluded. The experiment was conducted with the inspection circle on the left of the fixation point with 3 subjects, and on the right with the other 3.

Care was taken in exposing the figures not to introduce contours irrelevant to the experiment anywhere near the parts of the visual field upon which the figures were projected. All experiments were carried out in full daylight, though no direct sunlight was allowed to enter the room.

All the experiments were performed with binocular vision. Subjects with defective sight wore glasses to correct for it. None of the subjects used in experiment 4 had defective vision.

Subjects for the experiments were all volunteers. They were of both sexes and ages ranged from 18–32. All subjects were experimentally naïve. Some of the subjects were students, nurses or secretaries; many were foreign visitors to Oxford.

III

Results

The results of the first three experiments are presented in Table I.

It will be seen from this Table that altogether 5 subjects failed to comply with the instructions given them and refused to give a judgment to the effect that one circle looked bigger than the other. For this reason these subjects have not been included in the statistical analysis. It is sufficient for our purposes to show that where subjects do see a difference in the size of the test and comparison circles, the direction of that difference is correlated with the side upon which they have previously viewed the inspection figure. A Chi squared test was performed on each of these experiments

to discover whether the difference in the distribution of judgments where the inspection figure was on the left and where it was on the right was significant. The results of experiments (1) and (3) were significant at well beyond the 0·01 level of confidence, and those of experiment (2) were significant at the 0·001 level of confidence. The results of experiments (1) and (2) were in the direction which would be predicted if apparent size was a determining factor in this situation in producing the figural after-effect. In experiment (1), the test circle is seen as larger than the comparison circle presumably because in this experiment the inspection circle was of smaller apparent size than the test circle. In experiment (2), the test circle is seen as smaller than the comparison circle presumably because in this experiment the apparent size of the inspection circle is greater than that of the test circle. In both cases therefore the contours of the test circle moved away from the position of the contours of the inspection circle, if we consider the apparent sizes of the circles and not their retinal sizes which were equal.

TABLE I

THE RESULTS OF EXPERIMENTS 1, 2, AND 3

Experiment	*Position of Inspection Figure*	*Number of Subjects*	*Judgments on Test Figures*		X^2	*Level of Significance*	*Subjects judging Test Figures as equal*
			Left Bigger	*Right Bigger*			
1	Right	12	3	9	8·4	0·01	2
	Left	12	10	2			
2	Right	11	10	1	11·7	0·001	1
	Left	11	2	9			
3	Right	10	8	2	9·9	0·01	2
	Left	10	1	9			

The results of experiment (3) in which a similar situation was employed to that used by Hochberg and Bitterman (1951) are those that would be predicted on the basis of retinal size. As in Hochberg and Bitterman's experiment, retinal size and apparent size considered as factors in producing a figural after-effect are in conflict in this experiment, since moving the inspection figure nearer to the subject makes it retinally larger than the test figure though its apparent size remains smaller than that of the test figure. Thus the result of this experiment confirms the results already obtained by Hochberg and Bitterman and by Prentice viz. that in some situations retinal size may predominate over apparent size in producing a figural after-effect.

Subjects were asked at the end of the experiment to describe any changes that occurred in the inspection figure during the period of inspection. Most subjects reported changes of the type reported by Kohler and Wallach in their original monograph (1944). Changes reported (in order of frequency) were: (*a*) fading of different parts of the circle; (*b*) distortions of the circle's shape, e.g. the circle looked polygonal at times to some subjects, to others it appeared flattened out either in the vertical or horizontal directions; (*c*) a few subjects reported that the outlines of the circle

appeared to be moving in a circular direction; and (*d*) one subject reported that the circle appeared to be moving in the third dimension, i.e. to be alternately standing out from the wall to which it was attached and moving back to the wall. It is perhaps worth noting that a smaller proportion of the subjects who gave judgments at odds with the majority or who were unable to give a judgment at all saw such changes than of those subjects who gave judgments which tallied with the majority's judgments.

(4) Results for the fourth experiment are given in Table II.

TABLE II

THE RESULTS OF EXPERIMENT 4

	Inspection Circle on Left						*Inspection Circle on Right*								
Subject	1			2			3			4			5			6		
Distance from eyes in inches	L	=	R	L	=	R	L	=	R	L	=	R	L	=	R	L	=	R
47·6	4	0	6	10	0	0	2	2	6	2	6	2	2	4	4	2	6	2
51·6	0	0	10	1	3	6	3	3	4	4	6	0	4	2	4	4	6	0
55·6	0	0	10	0	8	2	2	5	3	6	4	0	10	0	0	6	4	0
57·6	0	0	10	0	2	8	0	3	7	8	2	0	10	0	0	8	2	0
59·6	0	0	10	0	2	8	0	1	9	10	0	0	9	1	0	7	3	0
63·6	0	3	7	1	0	9	1	2	7	7	3	0	10	0	0	10	0	0
67·6	0	1	9	0	0	10	0	1	9	10	0	0	10	0	0	10	0	0

The results of experiment (4) in general confirm our hypothesis. The basic situation is the same as in experiment (2). A figural after-effect which depended solely on apparent size would result in subjects seeing the test circle as smaller than the comparison circle since owing to constancy the apparent size of the inspection circle was larger than that of the test circle at all positions of the test circle. If retinal size alone were the determining factor the test circle should be seen as larger than the comparison circle at distances from the eye of less than 4 ft. 9·6 in., as smaller at distances beyond this. It is obvious that for all 6 subjects who completed the experiment the point at which the test figure has to be exposed in order that the comparison and test circles should appear equal is considerably nearer the eye than the point at which inspection and test figures are retinally equal (4 ft. 9·6 in.). The results are disappointing, however, for two reasons: (1) There are too many inconsistencies in the results of individual subjects for it to be worth attempting to determine the point for each subject at which the probabilities of obtaining a judgment of "left bigger" or "right bigger" are equal. (2) With several of the subjects the proportion of judgments which were in the opposite direction to the judgments which we would expect if apparent size were the determining factor became less as the experiment proceeded. This was particularly the case with subjects (1) and (6). It may be that subjects tended to repeat the judgment which they found to be most frequent, and owing to the design of the experiment this would in fact be the judgment based on apparent size. It should be noticed that none of the subjects realized that their judgments were being influenced by the position at which the test figure was exposed, though some of the subjects realized towards the end of the experiment that it was in fact being exposed at different distances from the eye.

IV

Discussion

The results of experiments (1) and (2) unambiguously support the hypothesis they were designed to test viz. that it is not merely retinal size that is important in determining figural after-effect but that the relative apparent sizes of test and inspection figures may also be a determining factor. The situation is, however, complicated by two further considerations:

Firstly, Kohler and Wallach found that when inspection and test figures were the same in size and when they were exposed at the same distance from the eye, either the test figure was seen as very slightly smaller than the comparison figure or the test and comparison figures were seen as of the same size. Now in experiment (2) the result might possibly be explained without introducing the factor of apparent size as part of the explanation since here the test figure is seen as smaller after inspection of the inspection figure whose contours fall on the same part of the retina as those of the test figure. The degree to which subject saw the test figure as smaller than the comparison figure in experiment (2) indicates that this is not, in fact, the correct explanation, since Kohler and Wallach found only a very small effect in this situation where test and inspection figures were in the same plane. Moreover, experiment (1) cannot be explained in this way since in this experiment the test figure is seen as larger than the comparison figure.

Secondly, we have assumed so far that figures of the same shape and angular size will, in fact, form identical images on the retina irrespective of the distances from the eye at which they are exposed. This is not the case. Duke-Elder (1932) states that as accommodation increases the size of the retinal image for a figure of given angular size increases. This fact, cannot, however, be used to explain our findings. Thus: (*a*) At the distances we are concerned with the effect is minute. (*b*) Any effect due to this factor would operate in the opposite direction to that due to apparent size, since while the retinal size of the near circle will be increased relatively to that of the further through the effects of accommodation, the effect that was demonstrated in experiments (1) and (2) was in the direction which would be predicted if the further circle is seen as larger than the nearer.

Experiment (3) is in effect a repeat of the experiment performed by Hochberg and Bitterman (1951). Our results show that with the degree of retinal separation of test and inspection figures produced by exposing them at a distance of 4 ft. and 12 ft. respectively, retinal size and not apparent size is the determining factor in producing the after-effect. This raises the question of what happens when the smaller figure is at intermediate distances. If both effects (i.e. that due to retinal size and that due to apparent size) are interacting we would predict that as the difference in retinal size is increased by moving the position of the smaller circle nearer to the eye and away from the point at which retinal size is approximately the same (where the angular distance of test and inspection figure is the same), there should be a decrease in judgments of the kind that would be expected on the basis of apparent size and an increase of the judgments of the kind that would be expected on the basis of retinal size. Moreover, since constancy is known to vary considerably from subject to subject we would predict that there might be considerable differences in the judgments given by different subjects with the smaller circle at varying distances, and in particular if both effects are occurring simultaneously we would expect that there would be a point for each subject where the two effects exactly counter-balance one another, and there is no figural after-effect, i.e. test and comparison figures are judged to be the same size.

The results of experiment (4) afford some confirmation for this prediction, but for the reasons given in the last section these results are not completely satisfactory. This experiment should be repeated using the method of adjustment instead of that of constant stimuli, since that would prevent subjects becoming fixated on a given response and it might be possible to determine with some accuracy for different subjects what distance the test figures would have to be exposed at in order that they should be seen as equal after the inspection periods.

As has already been indicated in the Introduction, the hypothesis that we regard as having been proved by the experiments we have undertaken enables us to explain the apparently conflicting results of earlier experiments and these experiments may be regarded as affording further confirmation of the hypothesis, namely that both retinal size and apparent size are of importance in determining the direction and amount of a given figural after-effect.

In summary then, experiments (1) and (2) prove that apparent size is one determining factor in producing figural after-effects. Experiments (3) and (4) indicate that retinal size is also of importance in determining figural after-effects and that the two factors may both be operative at the same time in producing a given after-effect.

The implications of this discovery are:—Firstly, perception is at least a two stage process otherwise it is impossible to explain how both factors can contribute to the figural after-effect at one and the same time. Secondly, whatever the mechanism behind the perception of apparent size and figural after-effects, the two mechanisms must be such that at some stage in the perception of apparent size the mechanism behind the production of figural after-effects can operate. It is hoped to present a second paper examining in detail the theoretical implications of the discovery that figural after-effects are influenced by apparent size, both for the mechanism behind figural after-effects and for that behind the constancies.

This research was made possible by a grant given me by the Institute of Experimental Psychology. My thanks are due to Professor G. Humphrey for affording me the facilities necessary for the conduct of the experiment, and to Mr. J. A. Deutsch for much helpful advice.

References

Duke-Elder, W. S. (1932). *Text Book of Opthalmology*. London. Vol. 1, p. 745.

Hochberg, J. E., and Bitterman, M. E. (1951). Figural after-effects as a function of the retinal size of the inspection figure. *Amer. J. Psychol.*, **64**, 99–102.

Kohler, W., and Emery, D. A. (1947). Figural affer-effects in the third dimension. *Amer. J. Psychol.*, **60**, 159–201.

Kohler, W., and Wallach, H. (1944). Figural after-effects. *Proc. Amer. Phil. Soc.*, **88**, 4, 269–357.

Osgood, C. E., and Heyer, A. W. Jr. (1952). A new interpretation of figural after-effects. *Psychol. Rev.*, **59**, 98–118.

Prentice, W. C. H. (1947). The relation of distance to the apparent size of figural after-effects. *Amer. J. Psychol.*, **60**, 617–23.

Prentice, W. C. H. (1950). On the size of the figural after-effect with varying distances. *Amer. J. Psychol.*, **63**, 589–93.

Proceedings of the Experimental Psychology Group, 1953

13TH–14TH APRIL, 1953.—Extended Meeting at Nottingham. *1st Session:* "The Origin of Adaptive Behaviour," by W. Ross Ashby (by invitation). Short communications by members of the Psychology Laboratory, University of Nottingham—L. J. Newson, Elizabeth Newson and G. Collins (all by invitation). *2nd Session:* "The Perceiving and Remembering of Instructional Television Programmes," by M. D. Vernon. "An Investigation of Intersensory Facilitation," by J. R. Symons (by invitation). *3rd Session:* "The Waning of an Instinctive Response," by R. A. Hinde (by invitation).

7TH–8TH JULY, 1953.—Extended Meeting at Cambridge. *1st Session:* "Some Maze Learning Experiments with Human Subjects," by J. von Wright. *2nd Session:* "Factors Affecting Immediate Memory," by J. Brown (by invitation). "The Use of Statistical Models in Psychology," by N. L. Webb (by invitation). *3rd Session:* Symposium on "Psychological Theories and Animal Behaviour," by H. Verplanck, J. S. Kennedy (both by invitation) and S. Koch. *4th Session:* Symposium (continued). Discussion opened by D. Russell Davis.

18TH–19TH SEPTEMBER, 1953.—Extended Meeting at Reading. *1st Session:* "Subception and Automisation in Relation to Theories of Inhibition," by D. E. Broadbent. Demonstration by R. C. Oldfield. *2nd Session:* "Experimenting with Instructional Films," by S. Laner (by invitation). "Subliminal Auditory Perception," by N. F. Dixon (by invitation). *3rd Session:* "Dogmatism and Rigidity as Determinants of Cognition," by Milton Rokeach (by invitation).

19TH DECEMBER, 1953.—7th Annual General Meeting at University College, London. *1st Session:* "Memory Systems in *Octopus vulgaris* Lamarck," by B. B. Boycott (by invitation). *2nd Session:* "A New Type of Behaviour Theory," by J. A. Deutsch (by invitation).

Members

BABINGTON SMITH, B., O.B.E., M.A., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
BARLOW, H., M.D., Department of Physiology, Downing Street, Cambridge.
BROADBENT, D. E., M.A., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
CANE, Miss V. R., M.A., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
CARPENTER, A., M.B., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
CHAMBERS, E. G., M.A., Psychological Laboratory, Downing Street, Cambridge.
CRAWFORD, Miss A., Ph.D., Department of Psychology, The University, Liverpool, 7.
CROWN, S., Ph.D., 81, Elgin Avenue, London, W.9.
DAVIS, D. RUSSELL, M.D. (*Editor*), Psychological Laboratory, Downing Street, Cambridge.
DEUTSCH, J. A., M.A., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
DREW, Prof. G. C., M.A., Department of Psychology, The University, Bristol.
FOSS, B. M., M.A., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
GREGORY, R. L., B.A., Psychological Laboratory, Downing Street, Cambridge.
GRINDLEY, G. C., M.A., Psychological Laboratory, Downing Street, Cambridge.
HALL, K. R. L., D.Phil., Barrow Hospital, Barrow Gurney, nr. Bristol.
HARPER, R., Ph.D., Department of Psychology, The University, Leeds, 2.
HARRIS, A. D., M.R.C.S., (*Hon. Secretary*), Psychological Laboratory, Downing Street, Cambridge.
HEIM, Miss A. W., Ph.D., Psychological Laboratory Downing Street, Cambridge.
HICK, W. E., M.D. (*Hon. Treasurer*), Psychological Laboratory, Downing Street, Cambridge.
HIMMELWEIT, Mrs. H. T., Ph.D., London School of Economics, Houghton Street, Aldwych, London, W.C.2.
HINDE, R. A., D.Phil., Ornithological Field Station, Madingley, Cambridge.
HOTOPH, N., M.A., London School of Economics, Houghton Street, Aldwych, London, W.C.2.

James, P. H. R., M.A., Department of Psychology, University College, Gower Street, London, W.C.1.
Jonckheere, A. R., B.Sc. (*Committee*), Department of Psychology, University College, Gower Street, London, W.C.1.
Joynson, R. B., M.A., (*Committee*), Department of Philosophy, The University, Nottingham.
Kay, H., Ph.D. (*Committee*), Institute of Experimental Psychology, 34, Banbury Road, Oxford.
Leonard, J. A., Ph.D., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
Mackworth, N. H., M.B., Ph.D., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
MacNeill, Miss F. E., Ph.D., Department of Psychology, The University, Manchester.
Oldfield, Prof. R. C., M.A., Department of Psychology, The University, Reading.
Pickford, R. W., D.Litt., Ph.D., Department of Psychology, The University, Glasgow, W.2.
Piercy, M. F., M.A., Department of Experimental Psychiatry, The University, Birmingham.
Poulton, E. C., M.B., Medical Research Council Unit for Applied Psychology, 15, Chaucer Road, Cambridge.
Russell, Prof. R. W., Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
Semeonoff, B., Ph.D., Department of Psychology, The University, Edinburgh.
Summerfield, A., B.Sc., Department of Psychology, University College, Gower Street, London, W.C.1.
Szafran, J., Ph.D., Psychological Laboratory, Downing Street, Cambridge.
Trist, E., O.B.E., M.A., Institute of Human Relations, 2, Beaumont Street, London, W.1.
Vernon, Miss M. D., D.Sc., Department of Psychology, The University, Reading.
Vince, Miss M. A., M.A., Psychological Laboratory, Downing Street, Cambridge.
Welford, A. T., M.A., Psychological Laboratory, Downing Street, Cambridge.
Whitfield, J. W., M.A. (*President*), Department of Psychology, University College, Gower Street, London, W.C.1.
Zangwill, Prof. O. L., M.A., Psychological Laboratory, Downing Street, Cambridge.

Visiting Foreign Members

Neff, Prof. W. D., Ph.D., Office of Naval Research, Keysign House, 429, Oxford Street, London, W.1.
Von Wright, J., M.A., Institute of Experimental Psychology, 34, Banbury Road, Oxford.

ZEITSCHRIFT FÜR EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

(*Journal of Experimental and Applied Psychology*)
(*Organ der Deutschen Gesellschaft für Psychologie*)

Herausgeber für Deutschland
Prof. Dr. J. von Allesch, Göttingen; Prof. Dr. Ph. Lersch, München

Herausgeber für Österreich
Prof. Dr. H. Rohracher, Wien

Herausgeber für die Schweiz
Prof. Dr. R. Meili, Bern

Redaktion
Dr. C. J. Hogrefe, Göttingen

Band I, Heft 4, 1953

H. SACHER, Zur testpsychologischen Querschnittsdiagnostik und statistischen Auswertstechnik.
J. L. PINILLOS u. J. C. BRENGELMANN, Bilderkennung als Persönlichkeitstest.
G. KLEINING, Die optischen Täuschungen.
H. KARL, Die Diagnostik der Antriebsstruktur im Farbpyramidentest.
H. HÄFNER, Über Wahrnehmungs- und Bedeutungsstrukturen und ihre Beziehungen zur emotionalen Einstellung.

All articles are provided with summaries in English and French.
The journal appears quarterly and contains 160 pages. Subscription is DM 64.— per volume and DM 18.— for single numbers.

VERLAG FÜR PSYCHOLOGIE · DR. C. J. HOGREFE · GÖTTINGEN
(Germany) · Postfach 414

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

This journal publishes original experimental investigations which contribute to the development of psychology as an experimental science.

MONTHLY, TWO VOLUMES PER YEAR

Subscription, $15.00 **Single Copy, $1.50**

Order from
AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 SIXTEENTH STREET, N.W.
WASHINGTON 6, D.C.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The Quarterly Journal of Experimental Psychology is published for the Experimental Psychology Group by W. Heffer & Sons Ltd., Petty Cury, Cambridge, England, to whom all business communications should be addressed.

The subscription is £1 10*s*. 0*d*. (post free) per volume of 4 numbers. Numbers appear quarterly. The cost of any single number is now 10*s*. 0*d*. (postage 2*d*.).

Undergraduate and graduates of less than three years' standing are offered a special rate of 15*s*. 0*d*. per volume. They should apply to the Editor, enclosing a remittance for 15*s*. 0*d*. and stating in which university department they have worked and what year they expect to graduate or did graduate.

All other subscriptions and orders should be placed with W. Heffer & Sons Ltd., or any bookseller.

★ ★ ★ ★

Papers submitted for publication and all other communications on Editorial matters should be addressed to "The Editor, *Quarterly Journal of Experimental Psychology*, Psychological Laboratory, Downing Street, Cambridge, England."

Only original papers are accepted for publication. The copyright in any material accepted for publication is retained by the *Journal*. Reproduction in whole or in part is forbidden, except (1) under terms of the *Fair Copying Declaration* of the Royal Society, or (2) when specific permission has been given by the Editor.

Intending contributors should pay attention to the conventions of the *Journal*, and especially to the form in which references are given. Manuscripts should be typewritten with double spacing and with wide margins. Diagrams, which should be kept to the minimum, must be submitted in a form suitable for reproduction without redrawing. They should be drawn in black ink on heavy paper; lettering should be inserted not on the diagram but on a tracing fixed over the diagram. Legends and tables should be typed on separate sheets of paper.

Contributors receive 20 copies of their paper free. If they require further copies they should apply to the Editor at the time of returning proofs.

PRINTED BY W. HEFFER & SONS LTD., CAMBRIDGE, ENGLAND